



Universidade de Aveiro Departamento de Biologia  
2011

**Helena Natividade  
Resende Costa**

**MARCADORES GENÉTICOS DE PREVISÃO DE  
FENÓTIPOS EM CONTEXTO FORENSE**



**Helena Natividade  
Resende Costa**

## **MARCADORES GENÉTICOS DE PREVISÃO DE FENÓTIPOS EM CONTEXTO FORENSE**

Dissertação apresentada à Universidade de Aveiro para cumprimento dos requisitos necessários à obtenção do grau de Mestre em Biologia Molecular e Celular, realizada sob a orientação científica do Prof. Doutor Luís Souto Miranda, Professor Auxiliar Convidado do Departamento de Biologia da Universidade de Aveiro.

Dedico este trabalho aos meus pais pelas oportunidades de crescimento em conhecimento e em caráter que me proporcionaram.

## **O júri**

Presidente

**Prof. Doutora Maria do Céu Gomes dos Santos**  
Professora Auxiliar Convidada do Departamento de Biologia da Universidade de Aveiro

**Prof. Doutor Francisco Manuel Andrade Corte Real Gonçalves**  
Professor Associado da Faculdade de Medicina da Universidade de Coimbra

**Prof. Doutor Luís Manuel Souto de Miranda**  
Professor Auxiliar Convidado do Departamento de Biologia da Universidade de Aveiro

## **Agradecimentos**

Agradeço, em primeiro lugar, ao Prof. Doutor Luís Souto, pela oportunidade de desenvolver o trabalho conducente à presente dissertação e pela orientação e acompanhamento prestados.

Agradeço também a um conjunto de pessoas que permitiram a obtenção de informação valiosa no âmbito da realização do presente trabalho, nomeadamente o Prof. Doutor Francisco Corte Real (Director da Delegação Centro do Instituto Nacional de Medicina Legal) o Dr. Bruno Antunes (Perito de Recolha de Vestígios no Local de Crime do Laboratório de Polícia Científica da Polícia Judiciária) a Mestre Inês Ferreira Leite (Professora Assistente de Direito Penal na Faculdade de Direito da Universidade de Lisboa), a Dra. Ana Luísa Zêzere (Procuradora do Ministério Público), a Prof. Doutora Helena Moniz (Professora Auxiliar da Faculdade de Direito da Universidade de Coimbra), o Prof. Doutor Guilherme de Oliveira (Professor Catedrático da Faculdade de Direito da Universidade de Coimbra) pelo encaminhamento das questões à Prof. Doutora Helena Moniz, e ao Dr. Sérgio Fernandes do Gabinete de Documentação e Direito Comparado.

Finalmente resta agradecer a todos os que, de alguma forma, me encorajaram a continuar a investir na minha formação e valorização pessoal, em particular a minha família e amigos.

## Palavras-chave

SNPs, AIMS, marcadores genéticos, previsão de fenótipos, genética forense, bioética, legislação.

## Resumo

A análise de DNA permite obter informações relativas à aparência dos dadores das amostras biológicas que poderão ser úteis no âmbito forense quando a identidade do dador não pode ser estabelecida com base nos métodos de identificação tradicionais. Essas informações podem ser obtidas de forma direta, analisando marcadores diretamente relacionados com a característica que se pretende prever, ou indireta, com base na determinação da ancestralidade biogeográfica do indivíduo e, a partir desse dado, prevendo-se o provável fenótipo manifestado com base no facto de que determinadas características apresentam variação interpopulacional.

Apesar de constituir uma ferramenta com o potencial de fornecer pistas importantes, a previsão de fenótipos em contexto forense com base na análise de marcadores genéticos apresenta várias limitações técnicas, sendo possível apenas prever com elevado nível de precisão os traços relacionados com a pigmentação e o género.

Os avanços técnicos, em particular o desenvolvimento de tecnologias de sequenciação de próxima geração, prometem, por um lado, possibilitar a realização de mais estudos de associação que permitirão conhecer as bases genéticas de outros traços físicos, e por outro, facilitar a aplicação de marcadores de previsão na rotina forense depois de validados.

Para além das limitações técnicas, a aplicação dos métodos de previsão levanta também questões de ordem bioética, nomeadamente a possibilidade de exacerbar ideais racistas, e também de ordem legal, em particular no que toca a potenciais ameaças ao direito à privacidade, identidade e imagem, discriminação étnica e racial e também as restrições relacionadas com o tipo de marcadores que podem ser usados no âmbito da análise forense.

Em Portugal os estudos populacionais indicam uma diversidade genética limitada, comparativamente a outros países. Para além disso, apesar de não haver legislação específica para este tipo de marcadores, a legislação existente é muito restritiva no que toca ao tipo de marcadores que podem ser usados no âmbito da análise forense de DNA e também no que toca aos direitos dos cidadãos. Assim, a aplicação destes marcadores no nosso país implica o prévio debate público e posterior criação de legislação específica.

## **Keywords**

SNPs, AIMs, genetic markers, predicting phenotypes, forensic genetics, bioethics, legislation.

## **Abstract**

DNA analysis provides information concerning the appearance of donors of biological samples that may be useful in forensics when the identity of the sample's donor can't be established based on traditional identification methods. This information can be obtained directly by analyzing markers directly related to the feature that is intended to predict, or indirect, based on the determination of the individual biogeographical ancestry, and from this data, predicting the probable phenotype expressed based on fact that certain characteristics show an interpopulation variation.

Although being a tool with the potential to provide important clues, predicting phenotypes in the forensic context based on the analysis of genetic markers has several technical limitations, and it's only possible to predict with high accuracy the pigmentation-related traits and gender.

Technical advances, particularly the development of technologies for next generation sequencing, promise, on the one hand, enable further studies of association that will allow knowing the genetic basis of other physical features, and by the other hand, to facilitate the application of the prediction markers in the forensic routine once validated.

In addition to the technical limitations, the application of the prediction methods also raises bioethical issues, including the possibility of exacerbating racist ideals, and also legal issues, in particular with regard to potential threats to the right to privacy, identity and image, ethnic and racial discrimination and also restrictions related with the type of markers that can be used in the forensic analysis.

In Portugal, population studies indicate a limited genetic diversity, comparatively to other countries. In addition although there is no specific legislation for this type of markers, the existing legislation is very restrictive with regard to the type of markers that can be used in the forensic analysis of DNA and also with regard to the citizens' rights. Thus, the application of these markers in our country requires the prior public debate and subsequent creation of specific legislation.





# ÍNDICE

ÍNDICE DE FIGURAS .....	V
ÍNDICE DE TABELAS .....	IX
SIGLAS .....	XI
 <b>INTRODUÇÃO .....</b>	 <b>1</b>
<b>OBJETIVOS .....</b>	<b>2</b>
 <b><u>CAPÍTULO 1 .....</u></b>	 <b><u>3</u></b>
 <b><u>1. MARCADORES GENÉTICOS COM APLICAÇÃO FORENSE .....</u></b>	 <b><u>3</u></b>
<b>1.1. CARACTERÍSTICAS GERAIS E PADRÕES DE HERANÇA DOS MARCADORES GENÉTICOS .....</b>	<b>5</b>
1.1.1. MARCADORES AUTOSSÔMICOS .....	6
1.1.2. MARCADORES DO CROMOSSOMA X .....	9
1.1.3. MARCADORES MITOCONDRIAIS .....	11
1.1.4. MARCADORES DO CROMOSSOMA Y .....	13
<b>1.2. POLIMORFISMOS DE DNA QUE SÃO USADOS COMO MARCADORES GENÉTICOS .....</b>	<b>14</b>
1.2.1. SHORT TANDEM REPEATS (STRs) .....	15
1.2.1.1. A ORIGEM DE UM STR .....	17
1.2.1.2. CARACTERÍSTICAS DOS MARCADORES DE STR .....	18
1.2.2. SINGLE NUCLEOTIDE POLYMORPHISMS (SNPs) .....	20
1.2.2.1. DESEQUILÍBRIO DE LIGAÇÃO E HAPLÓTIPOS .....	23
1.2.2.2. CLASSIFICAÇÃO DOS SNPs EM CONTEXTO FORENSE .....	25
1.2.3. SNPs vs STRs .....	25
<b>1.3. MARCADORES GENÉTICOS QUE REFLETEM A DIVERSIDADE INTERPOPULACIONAL .....</b>	<b>27</b>
1.3.2. ORIGEM E EVOLUÇÃO DA ESPÉCIE HUMANA .....	32
1.3.3. MARCADORES INFORMATIVOS DE ANCESTRALIDADE .....	36
1.3.3.1. HAPLOGRUPOS DA NRY E DO MTDNA .....	39

**CAPÍTULO 2 .....45**

**2. METODOLOGIAS DE ANÁLISE DE SNPS.....45**

**2.1. REVERSE DOT BLOTS .....48**

**2.2. PCR QUANTITATIVA EM TEMPO REAL.....49**

**2.3. PIROSSEQUENCIAÇÃO .....51**

**2.4. MINISSEQUENCIAÇÃO.....52**

**2.5. ESPECTROMETRIA DE MASSA .....54**

**2.6. MICROARRAYS DE DNA.....56**

**2.7. MÉTODOS DE SEQUENCIAÇÃO DE PRÓXIMA GERAÇÃO E PERSPETIVAS FUTURAS .....58**

**CAPÍTULO 3 .....61**

**3. APLICAÇÕES PRÁTICAS DOS MARCADORES GENÉTICOS EM CONTEXTO FORENSE**

**.....61**

**3.1. PREVISÃO DE FENÓTIPOS – MÉTODO DIRETO .....61**

**3.1.1. GÊNERO OU SEXO .....63**

**3.1.2. TRAÇOS RELACIONADOS COM A PIGMENTAÇÃO .....64**

**3.1.2.1. COR DOS OLHOS .....67**

**3.1.2.2. COR DO CABELO .....71**

**3.1.2.3. TOM DE PELE.....74**

**3.1.2.4. ALTURA.....76**

**3.1.2.5. TRAÇOS FACIAIS.....78**

**3.1.2.6. IDADE .....79**

**3.2. ANCESTRALIDADE BIOGEOGRÁFICA .....83**

**3.2.1. ANCESTRALIDADE INDIVIDUAL .....83**

**3.2.2. PAINÉIS DE AIMS.....86**

**3.2.2.1. PAINÉIS DE AIMS COM APLICAÇÃO FORENSE.....90**

**3.2.3. PREVISÃO DE FENÓTIPOS A PARTIR DA ANCESTRALIDADE BIOGEOGRÁFICA – MÉTODO INDIRETO.....92**

**3.2.3.1. TRAÇOS RELACIONADOS COM A PIGMENTAÇÃO .....93**

**3.2.3.2. MORFOLOGIA DO CABELO .....100**

**3.2.3.3. ÍNDICE DE MASSA CORPORAL .....102**

**3.2.3.4. TRAÇOS FACIAIS.....103**

**3.2.4. FERRAMENTAS DE PREVISÃO INDIRETA DE FENÓTIPOS .....103**

<b>3.3. CASOS PRÁTICOS .....</b>	<b>106</b>
<b>3.4. A POPULAÇÃO PORTUGUESA .....</b>	<b>108</b>
3.4.1. MARCADORES AUTOSSÓMICOS.....	108
3.4.2. MARCADORES DO CROMOSSOMA X .....	110
3.4.3. MARCADORES DO CROMOSSOMA Y .....	111
3.4.4. MARCADORES DO MTDNA.....	113
 <b><u>CAPÍTULO 4 .....</u></b>	 <b><u>117</u></b>
 <b><u>4. QUESTÕES BIOÉTICAS E LEGAIS .....</u></b>	 <b><u>117</u></b>
4.1. QUESTÕES BIOÉTICAS.....	117
4.4. LEGISLAÇÃO INTERNACIONAL .....	124
4.5. EM PORTUGAL.....	125
 <b><u>DISCUSSÃO, CONCLUSÕES E PERSPETIVAS FUTURAS.....</u></b>	 <b><u>129</u></b>
 <b><u>REFERÊNCIAS BIBLIOGRÁFICAS.....</u></b>	 <b><u>137</u></b>
 <b><u>ANEXOS .....</u></b>	 <b><u>155</u></b>



## ÍNDICE DE FIGURAS

Figura 1 – Porções das diferentes classes de sequências nos genoma nuclear e mitocondrial.....	5
Figura 2 – Cromossomas humanos.....	6
Figura 3 – Esquema ilustrativo dos fenómenos envolvidos na hereditariedade genética, nomeadamente a gametogénese e a fecundação.....	7
Figura 4 – Mapas genéticos de três irmãos Jacob, Jolinard e Daniel O’Neil .....	8
Figura 5 – Esquema ilustrativo do modo de herança de um par de autossomas .....	9
Figura 6 – Esquemas ilustrativos do modo de transmissão do cromossoma X. ....	10
Figura 7 – Representação esquemática do DNA mitocondrial humano. ....	11
Figura 8 – Esquema representativo da região de controlo (D-loop) onde se encontram discriminadas as regiões hipervariáveis (1 e 2). ....	12
Figura 9 – Árvore genealógica onde se encontra representado o modo de transmissão do mtDNA. ....	13
Figura 10 - Árvore genealógica onde se encontra representado o modo de transmissão do cromossoma Y. ....	14
Figura 11 – Esquema representativo da origem de um microssatélite documentada através da comparação da sequência de nucleótidos observada em diferentes espécies. ....	17
Figura 12 - Esquema ilustrativo do processo de derrapagem de replicação .....	18
Figura 13 – Esquema representativo da forma mais comum de SNP: substituição de um nucleótido .....	20
Figura 14 – Esquema ilustrativo das possíveis consequências da ocorrência de SNPs em diversos locais do genoma.....	21
Figura 15 – Esquema representativo da distribuição genómica da densidade de SNPs nos cromossomas humanos 1 a 22.....	23
Figura 16 – Esquema ilustrativo da organização do material genético em haploblocos em cromossomas de dois indivíduos diferentes .....	24
Figura 17 – Representação esquemática do efeito “gargalo de garrafa”.....	29
Figura 18 – Gráfico de barras ilustrativo das proporções de mistura de populações ancestrais (representadas pelas diferentes cores) observadas em indivíduos de diferentes países.....	30
Figura 19 – Mapa representativo da migração do <i>Homo sapiens</i> moderno onde se encontram representadas as datas de chegada nas respetivas regiões de acordo com dados genéticos.....	34
Figura 20 – Esquema ilustrativo de dois modelos da origem dos seres humanos modernos – o modelo de evolução multirregional e o modelo “Arca de Noé” .....	35
Figura 21 – Mapa representativo da diversidade genética .....	36
Figura 22 – Esquemas representativos da filogenia do cromossoma Y .....	41
Figura 23 – Árvore filogenética simplificada onde se encontram os diversos haplogrupos do mtDNA representados pelas respetivas letras.....	42
Figura 24 – Árvore genealógica na qual estão representados os diferentes modos de herança dos marcadores autossómicos, NRY e mtDNA .....	43
Figura 25 – Esquemas ilustrativos de dois sistemas de incorporação de fluoróforos em reações PCR em tempo real .....	49

## Índice de Figuras

Figura 26 – Gráfico de amplificação de PCR em tempo real .....	50
Figura 27 – Esquema representativo do princípio do método de pirosequenciação.....	52
Figura 28 – Esquema representativo da minissequenciação.....	53
Figura 29 – Esquema ilustrativo do funcionamento de um espectrómetro de massa.....	54
Figura 30 – Representação esquemática da produção de partículas com carga e posterior representação gráfica dos resultados obtidos através de dois métodos de espectrometria de massa .....	55
Figura 31 – Representação esquemática do princípio de hibridização de microarray .....	57
Figura 32 – Estrutura molecular dos pigmentos melânicos.....	65
Figura 33 – Representação esquemática de um melanócito e vias de síntese de melanina.....	66
Figura 34 – Representação esquemática da localização de um melanócito e outras estruturas intervenientes no processo de pigmentação da pele.....	67
Figura 35 – Figura ilustrativa da variação da cor da íris desde o azul ao castanho escuro passando por tons intermédios (cinza, verde, avelã e castanho claro) .....	70
Figura 36 – Criança com fissura labial ou “lábio leporino” .....	78
Figura 37 – Detecção da deleção 4.977 bp do mtDNA.....	80
Figura 38 – Esquema ilustrativo da localização dos telómeros num cromossoma e sequência repetitiva que os constitui .....	81
Figura 39 – Estrutura populacional estimada recorrendo à análise estatística com o STRUCTURE dos dados obtidos através da aplicação de um painel de 377 STRs autossomais .....	84
Figura 40 – Mapa representativo da variabilidade genética interpopulacional.....	85
Figura 41 – Mapas representativos da distribuição geográfica do nível médio de heterozigotia e LD em 23 sub-populações Europeias.....	86
Figura 42 – Representação esquemática da comparação da estrutura populacional de uma amostra em cinco grupos (k=5) quando analisada com marcadores mais e menos informativos recorrendo ao STRUCTURE .....	89
Figura 43 – Mapa representativo da caracterização da origem biogeográfica obtida com a aplicação dos kits Ancestry by DNA 2.5 e EurasianDNA 1.0 .....	90
Figura 44 – Representação esquemática dos dados genotípicos obtidos por (NOVEMBRE et al., 2008) recorrendo a uma abordagem de PCA em dois eixos.....	92
Figura 45 – Fotografia de vários antebraços representando a diversidade fenotípica do tom de pele do mais escuro ao mais claro.....	94
Figura 46 – Níveis médios anuais de intensidade da radiação ultravioleta indicados por gradação de cor, sendo tanto mais elevados quanto mais clara a cor representada.....	95
Figura 47 – Distribuição do tom de pele nos seres humanos.....	95
Figura 48 – Correlação entre o Índice de Melanina e a Ancestralidade Genómica Nativo Americana, Oeste-Africana e Europeia determinada com base na aplicação de um painel de 171 AIMs descrito por (FRUDAKIS, 2008) a uma amostra de mulheres Porto Riquenhas .....	96
Figura 49 – Distribuição geográfica das percentagens de frequência de olhos claros na Europa e norte de África.....	97
Figura 50 – Fotografias de indivíduos que surgiram como resultado de uma pesquisa feita na base de dados EURO 1.0 correspondendo ao critério de pesquisa “NOR $\geq$ 75%” e “NOR $\leq$ 30%” (NOR – ancestralidade genómica Norte-Europeia) .....	98

Figura 51 – Frequências dos alelos A e T do SNP rs 11803731 do gene TCHH com base no <i>Human Genome Diversity Project</i> .....	101
Figura 52 – Exemplos ilustrativos da regra de Bergman.....	102
Figura 53 – Utilização de base de dados para estabelecimento de relações entre diferentes proporções de origem biogeográfica Africana e características fenotípicas.....	104
Figura 54 – Exemplo de pesquisa e nove resultados obtidos na base de dados referida no trabalho de FRUDAKIS (2008) recorrendo a uma estimativa de mistura individual de ancestralidade genómica para um indivíduo Caucasiano do sexo feminino.....	105
Figura 55 – Atentado de 11 de março de 2004 em Madrid.....	107
Figura 56 – Alinhamento de suspeitos para identificação.....	118
Figura 57 – Mapa no qual se encontra representada a distribuição geográfica das raças humanas definidas por Carleton Coon .....	122
Figura 58 – Dois exemplos do passado de situações extremas que constituíram consequências do domínio de ideais racistas: a subjugação de seres humanos – escravatura e o extermínio de um povo – Holocausto .....	123





## ÍNDICE DE TABELAS

Tabela 1 – Descrição e exemplos de fatores que influenciam as características fenotípicas complexas .....	4
Tabela 2 – Características dos genomas nuclear e mitocondrial.....	6
Tabela 3 – Teor de STRs do genoma humano .....	16
Tabela 4 – Principais características que distinguem os STRs e os SNPs no âmbito da análise forense de DNA.....	27
Tabela 5 – Resumo das principais características das forças evolutivas relativamente aos efeitos das mesmas ao nível da diferenciação interpulacional e da extensão de envolvimento do genoma....	32
Tabela 6 – Principais características relativas aos modelos de evolução dos humanos modernos. ..	35
Tabela 7 – Comparações de valores de $F_{ST}$ baseada em 369 pares de SNPs de populações ao longo das regiões major .....	38
Tabela 8 – Tabela onde estão representados as principais metodologias utilizados na genotipagem de SNPs no âmbito forense e respetivas características relativamente à discriminação alélica, formato da reação e mecanismos de deteção.....	47
Tabela 9 – Relação entre a cor da íris e a concentração de eumelanina e feomelanina presentes nessa mesma estrutura. ....	68
Tabela 10 – Genes e respetivos SNPs presentes nos três mais recentes conjuntos de marcadores que permitem a previsão da cor dos olhos e cabelo e tom de pele. ....	74
Tabela 11 – Resumo dos dados relativos à percentagem da variação fenotípica associada ao número de marcadores utilizados de trabalhos relativos aos marcadores de DNA relacionados com a altura do adulto.....	77
Tabela 12 – Lista de referências bibliográficas correspondentes a publicações de painéis de marcadores que permitem a determinação da origem biogeográfica de acordo com diferentes níveis de distinção.....	87
Tabela 13 – Conjunto de 19 trabalhos relativos à pigmentação humana classificados de acordo com o tipo de estudo – previsão ou associação entre SNPs e fenótipo e relação entre SNPs de genes relacionados com a pigmentação e origem biogeográfica.....	99
Tabela 14 – Principais sistemas de classificação racial com base em TISHKOFF et al. (2004).....	121
Tabela A1 – Dados de 19 estudos relativos às bases genéticas dos traços de pigmentação (cor dos olhos, cor do cabelo e tom de pele) na qual se encontram descritas as associações estabelecidas entre diversos SNPs de vários genes e as referidas características bem como as populações nas quais essas relações foram estabelecidas.....	156
Tabela A2 – Estudos relativos a frequências alélicas de diversos marcadores que incluem amostras de populações Portuguesas.....	167
Tabela A2 – Estudos relativos às frequências alélicas de diversas populações Portuguesas classificados de acordo com as populações amostradas.....	169



## SIGLAS

**ABI** – Applied Biosystems

**ADN** – Ácido Desoxirribonucleico

**AIM** – Ancestry Informative Marker

**ALFRED** – *the Allele Frequency Database*

**CNPD** – Comissão Nacional de Proteção de Dados

**CODIS** – *The Combined DNA Index System*

**D-loop** – Displacement loop

**DNA** – *Deoxyribonucleic Acid* (mesmo que ADN)

**ESI** – *Electrospray Ionization*

**FBI** – *Federal Bureau of Investigation*

**Flow-FISH** – *Flow Fluorescent in-situ Hybridization*

**FRET** – *Fluorescence Resonance Energy Transfer*

**GIANT** – *Genetic Investigation of Anthropometric Traits*

**GWAS** – *Genome Wide Association Studies*

**HGDP-CEPH** – *Human Genome Diversity Cell Line Panel*

**HLA** – *Human Leukocyte Antigen*

**HV** – *Hypervariable Region*

**LD** – *Linkage Disequilibrium*

**M** – Índice de Melanina

**MALDI** – *Matrix-Assisted Laser Desorption/Ionization*

**MHC** – *Major Histocompatibility Complex*

**mRNA** – *Messenger Ribonucleic Acid*

**MSY** – *Male-Specific Region of the Y Chromosome*

**mtDNA** – *Mitochondrial Deoxyribonucleic Acid*

**NCBI** – *National Centre for Biotechnology Information*

**NGS** – *Next-Generation Sequencing*

**NRPY** – *Non-Recombining Portion of the Y Chromosome*

**NRY** – *Non-Recombining Region of the Y Chromosome*

**OMIM** – *Online Mendelian Inheritance in Man*

**PCA** – *Principal Component Analysis*

**PCR** – *Polymerase Chain Reaction*

**Q-FISH** – *Quantitative Fluorescent in-situ Hybridization*

**Q-PCR** – *Quantitative Polymerase Chain Reaction*

**rCRS** – *Revised Cambridge Reference Sequence*

**RFLP** – *Restriction Fragment Length Polymorphism*

**RNA** – *Ribonucleic Acid*

**rRNA** – *Ribosomal Ribonucleic Acid*

**SIDA** – *Síndrome da Imunodeficiência Adquirida*

**sjTRECs** – *Signal Joint T-cell Receptors Excision Circles*

**SNP** – *Single Nucleotide Polymorphism*

**STRs** – *Short Tandem Repeats*

**TCR** – *T-cell receptor*

**TMB** – *Tetrametilbenzidina*

**TOF** – *Time Of Light*

**tRNA** – *Transfer Ribonucleic Acid*

**UV** – *Radiação Ultra Violeta*

**UVA** – *Radiação Ultra Violeta A*

**UVB** – *Radiação Ultra Violeta B*

**VNTR** – *Variable Number of Tandem Repeat*

**YCC** – *Y Chromosome Consortium*

**YHRD** – *Y Chromosome Haplotype Reference Database*



## INTRODUÇÃO

Quando uma amostra problema (amostra sob investigação, cuja identificação se pretende estabelecer) dá entrada num serviço de genética forense é determinado o perfil de identificação dessa mesma amostra (amostra problema) que é posteriormente comparado com outros perfis (de amostras referência – amostras relativamente às quais se conhece a identidade dos dadores – e/ou os perfis que constam na base de dados nacional) com o objetivo de se proceder à identificação do dador da mesma (Lei n.º 5/2008 de 12 de Fevereiro). No entanto, muitos são os casos em que a identificação não é estabelecida por não haver correspondência entre o perfil da amostra problema e qualquer perfil das amostras referência e/ou perfis da base de dados (KAYSER et al., 2009).

A informação contida no material genético das amostras biológicas não se limita, no entanto, à identificação. De facto, muitas características físicas (cor dos olhos, do cabelo e da pele (REES, 2003, SULEM et al., 2007), altura (LANGO ALLEN et al., 2010), entre outras) e até comportamentais (comportamento suicida (BONDY et al., 2006, ÖZALP, 2009), comportamento agressivo (CRAIG et al., 2009), comportamento criminal (ALPER, 1998) dependência do álcool (WANG et al., 2011, ZUO et al., 2011), entre outros) têm por base fatores genéticos que, caso sejam conhecidos, podem ser utilizados para traçar um perfil do dador de uma amostra cuja identidade não foi possível estabelecer pela via da identificação genética. Para além disso, uma vez que determinadas características físicas apresentam um padrão geográfico de variação (sendo o tom de pele o exemplo mais evidente) a previsão dessas mesmas características pode ser feita de forma indireta com base na informação relativa à origem biogeográfica do indivíduo (FRUDAKIS, 2008, KAYSER et al., 2009).

A previsão de características físicas do dador de uma amostra a partir da análise do DNA, ao contrário do perfil de identificação, não tem por objetivo identificar inequivocamente o dador da amostra, mas sim limitar o número de suspeitos a considerar no âmbito de uma dada investigação criminal (FRUDAKIS, 2008, KAYSER et al., 2009).

Em Portugal não são realizadas análises de DNA no contexto de previsão de características fenotípicas, segundo as informações obtidas junto do Instituto Nacional de

## Introdução

Medicina Legal e Laboratório da Polícia Científica. O Prof. Doutor Francisco Corte Real<sup>1</sup> mostrou-se bastante otimista relativamente às perspectivas de utilização destes marcadores na rotina forense, afirmando: “com o conhecimento cada vez maior das potencialidades deste tipo de marcadores, entendo que entrarão na rotina em breve”, ressaltando, no entanto que considera, nesta altura, precoce a sua utilização devido à “taxa de sucesso actualmente possível”. Por seu turno, o Dr. Bruno Antunes<sup>2</sup> refere ainda que a não utilização destes métodos de análise se deve por um lado, à inexistência de meios técnicos que possibilitem tais análises e, por outro, a limitações legais.

O presente trabalho visa, então, constituir uma revisão do conhecimento relativo à previsão de características fenotípicas úteis no contexto da investigação forense e sua aplicação em Portugal e noutros países, focando quatro aspetos fundamentais: conhecimentos teóricos relacionados com os marcadores genéticos de previsão, metodologias de análise desses mesmos marcadores (particularmente dos SNPs por serem os mais frequentemente aplicados na previsão direta e indireta), aplicações práticas dos marcadores de previsão em contexto forense e, finalmente, as questões bioéticas e legais subjacentes à previsão de fenótipos em contexto forense.

Nesse sentido, para além da pesquisa bibliográfica nas áreas da genética forense, antropologia molecular, bioética e direito foram também estabelecidos contactos com instituições de interesse no âmbito do trabalho, nomeadamente o Instituto Nacional de Medicina Legal, o Laboratório da Polícia Científica da Polícia Judiciária, o Ministério Público, a Comissão Nacional de Proteção de Dados e o Gabinete de Documentação e Direito Comparado.

## OBJETIVOS

O presente trabalho tem como objetivos: (1) descrever os aspetos teóricos, técnicos, bioéticos e legais subjacentes à previsão de ancestralidade e fenótipos a partir do DNA; (2) aferir a importância e a aplicabilidade do conhecimento relativo à previsão de ancestralidade e fenótipos na área forense em Portugal e noutros países.

---

<sup>1</sup> Prof. Doutor Francisco Corte Real, Director da Delegação Centro do Instituto Nacional de Medicina Legal – comunicação pessoal.

<sup>2</sup> Dr. Bruno Antunes, Perito de Recolha de Vestígios no Local de Crime do Laboratório de Polícia Científica da Polícia Judiciária – comunicação pessoal.

# CAPÍTULO 1

## 1. MARCADORES GENÉTICOS COM APLICAÇÃO FORENSE

As características observadas numa célula ou num organismo são o produto direto de atuação das proteínas que, por sua vez, são codificadas pelo DNA. Essas mesmas características correspondem ao fenótipo (STRACHAN et al., 2010). O conceito de fenótipo é, porém, um pouco mais amplo, incluindo não só as características bioquímicas ou fisiológicas, mas também a morfologia, o desenvolvimento e o comportamento do organismo. O fenótipo diz, então, respeito às características observáveis numa célula ou num organismo e resulta da expressão dos genes do mesmo e também da influência de fatores ambientais.

As características fenotípicas podem ser simples ou complexas, consoante sejam determinadas pela expressão de apenas um gene (características simples) ou sejam o resultado de vários fatores (PULKER et al., 2007), que se encontram descritos na Tabela 1 (características complexas).

A maioria das características físicas que definem a aparência dos seres humanos são características complexas, sendo, por isso, afetadas por um ou vários fatores (anteriormente referidos) o que dificulta a identificação dos fatores genéticos que estão na base dessas mesmas características (PULKER et al., 2007). A identificação de fatores genéticos relacionados com as características fenotípicas implica o estudo de marcadores genéticos.

**Tabela 1 – Descrição e exemplos de fatores que influenciam as características fenotípicas complexas (os fatores enumerados têm por base o trabalho de PULKER et al. (2007) e as descrições baseiam-se nas publicações de LENZ (1970) e STRACHAN et al (2010)).**

<b>Fator</b>	<b>Descrição</b>
Heterogeneidade alélica	Observa-se quando diferentes alelos do mesmo locus ou gene originam a mesma característica
Heterogeneidade de locus	Observa-se quando diferentes alelos de diferentes loci ou genes originam a mesma característica
Penetrância	Fração de indivíduos que manifestam uma dada característica quando possuem um determinado gene
Expressividade	Grau de efeito produzido pelo gene ou grau de manifestação clínica
Interação entre genes ou poligenia	Observa-se quando uma dada característica resulta da ação combinada de vários genes.
Interação entre genes e ambiente	Corresponde à forma como a expressão de um gene é condicionada por fatores ambientais
Fenocópia	Característica observada num organismo que, apesar de semelhante à determinada geneticamente, é determinada por fatores ambientais, não sendo transmitida à descendência

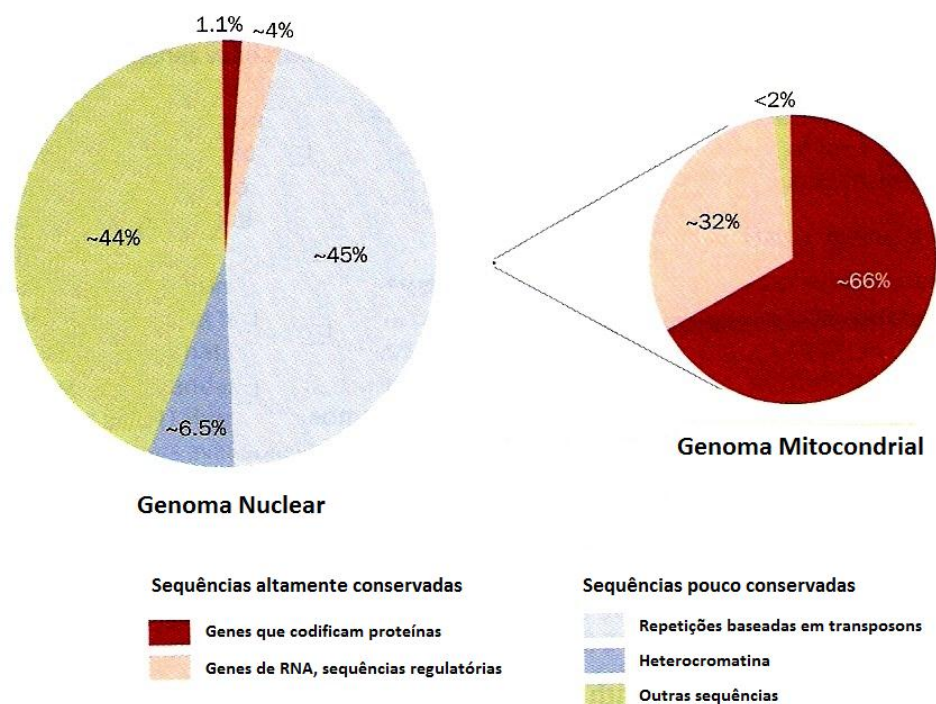
Os marcadores genéticos são segmentos de DNA cuja sequência e posição no genoma são conhecidas. A análise de marcadores genéticos possibilita a realização de estudos em diversas áreas, nomeadamente na área forense (na identificação em casos de crimes, catástrofes ou pessoas desaparecidas, acidentes militares e testes de filiação e parentesco: questões de imigração e reivindicação de heranças) e na área biomédica (autenticação de linhas celulares humanas, monitorização de transplantes, deteção de tumores cancerígenos e mapeamento de doenças genéticas). Existe uma área de investigação – o estudo da diversidade das populações humanas – que tem interesse tanto do ponto de vista forense como clínico, visto ser uma área que permite obter informações valiosas no contexto da previsão da ancestralidade (com aplicações na área forense) e também no contexto de risco de doença associado com a origem biogeográfica. O conhecimento relativo à diversidade das populações tem também bastante relevância no contexto dos estudos acerca da evolução da espécie humana (BUTLER et al.).



Segue-se a descrição dos diversos tipos de marcadores quanto às características gerais e padrões de herança e quanto à sua natureza (polimorfismos de DNA).

### 1.1. CARACTERÍSTICAS GERAIS E PADRÕES DE HERANÇA DOS MARCADORES GENÉTICOS

O genoma humano subdivide-se em genoma nuclear, com cerca de 3,1 Gb que compreende os autossomas e cromossomas sexuais X e Y, e genoma mitocondrial com 16,6 kb (STRACHAN et al., 2010). Apesar do número de genes do genoma nuclear (cerca de 26.000) ser maior do que o observado no genoma mitocondrial (37), a percentagem que este tipo de sequência representa é francamente menor no genoma nuclear do que no genoma mitocondrial (STRACHAN et al., 2010) como é visível na Figura 1. As principais características que distinguem os genomas nucleares e mitocondrial encontram-se sumariadas na Tabela 2.



**Figura 1 – Porções das diferentes classes de sequências nos genoma nuclear e mitocondrial. O pequeno ponto ao centro da figura representa o equivalente a 25 genomas mitocondriais à escala do genoma nuclear à esquerda. Outra diferença bem patente reside na predominância de sequências altamente conservadas no genoma mitocondrial que contrasta com o observado no genoma nuclear. Figura adaptada de STRACHAN et al. (2010).**

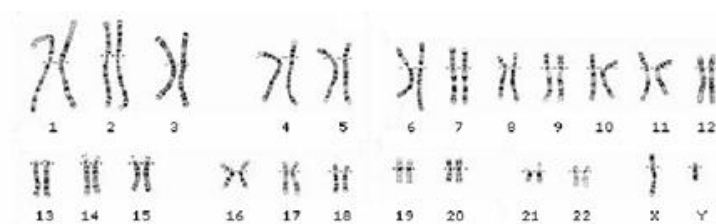
**Tabela 2 – Características dos genomas nuclear e mitocondrial. Adaptado de STRACHAN et al. (2010).**

Características	Genoma Nuclear	Genoma Mitocondrial
Proteínas associadas	Vários tipos (incluindo histonas)	Praticamente livre de proteínas
Número de genes que codificam proteínas	~21.000	13
Número de Genes de RNA	> 6.000 (incerto)	24
Densidade génica	~1/120 kb (incerto)	1/0,45 kb
DNA repetitivo	>50% do genoma	Muito pouco
Intrões	Presentes na maioria dos genes	Ausentes
Recombinação*	Pelo menos uma vez por cada par de cromossomas homólogos na meiose	Não ocorre
Herança*	Mendeliana para o cromossoma X e autossomas; paterna para o cromossoma Y	Materna

\*Assuntos a desenvolver adiante.

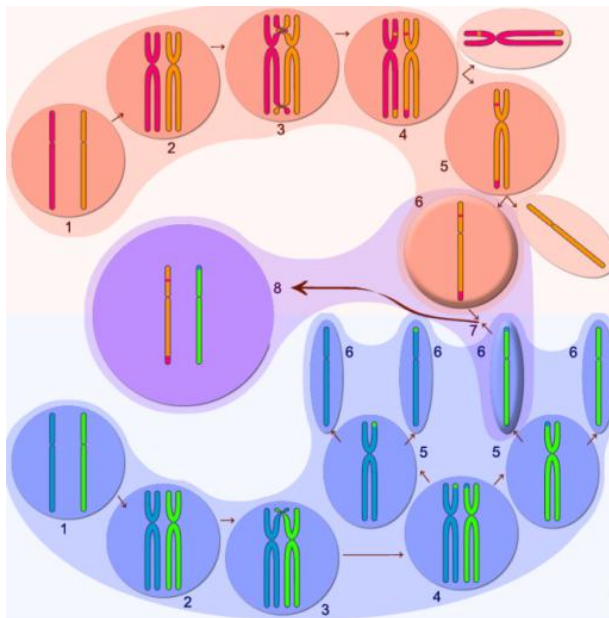
### 1.1.1. MARCADORES AUTOSSÓMICOS

O DNA nuclear encontra-se organizado em 23 pares de cromossomas (Figura 2), sendo que 22 desses pares são autossomas e o par restante corresponde aos cromossomas sexuais (cromossoma X e cromossoma Y). Nos indivíduos do sexo feminino, a combinação de cromossomas sexuais observada é XX (dois cromossomas X) enquanto nos indivíduos do sexo masculino a combinação observada é XY (um cromossoma X e um cromossoma Y).



**Figura 2 – Cromossomas humanos. Fotografia de um cariótipo humano (conjunto ordenado de cromossomas) na qual se podem observar os 22 pares de autossomas e os cromossomas sexuais X e Y; adaptado de NATIONAL GEOGRAPHIC SOCIETY.**

Os autossomos (cromossomas 1 a 22) constituem a maior porção do património genético de cada ser humano. Um dos elementos de cada par de cromossomas é herdado do pai e o outro da mãe, sendo que este fenómeno é apenas possível graças ao processo de meiose que ocorre aquando da gametogénese (Figura 3). Este processo tem início numa célula diplóide e envolve a ocorrência de uma replicação do DNA seguida de duas divisões, o que se traduz num resultado final de quatro células haplóides, ou seja, células que contêm metade da quantidade de material genético da célula inicial. Quando se dá a fecundação entre os gâmetas (células haplóides) forma-se uma célula diplóide – ovo ou zigoto – cujo material genético provém dos dois progenitores em partes iguais (no que diz respeito aos autossomas).



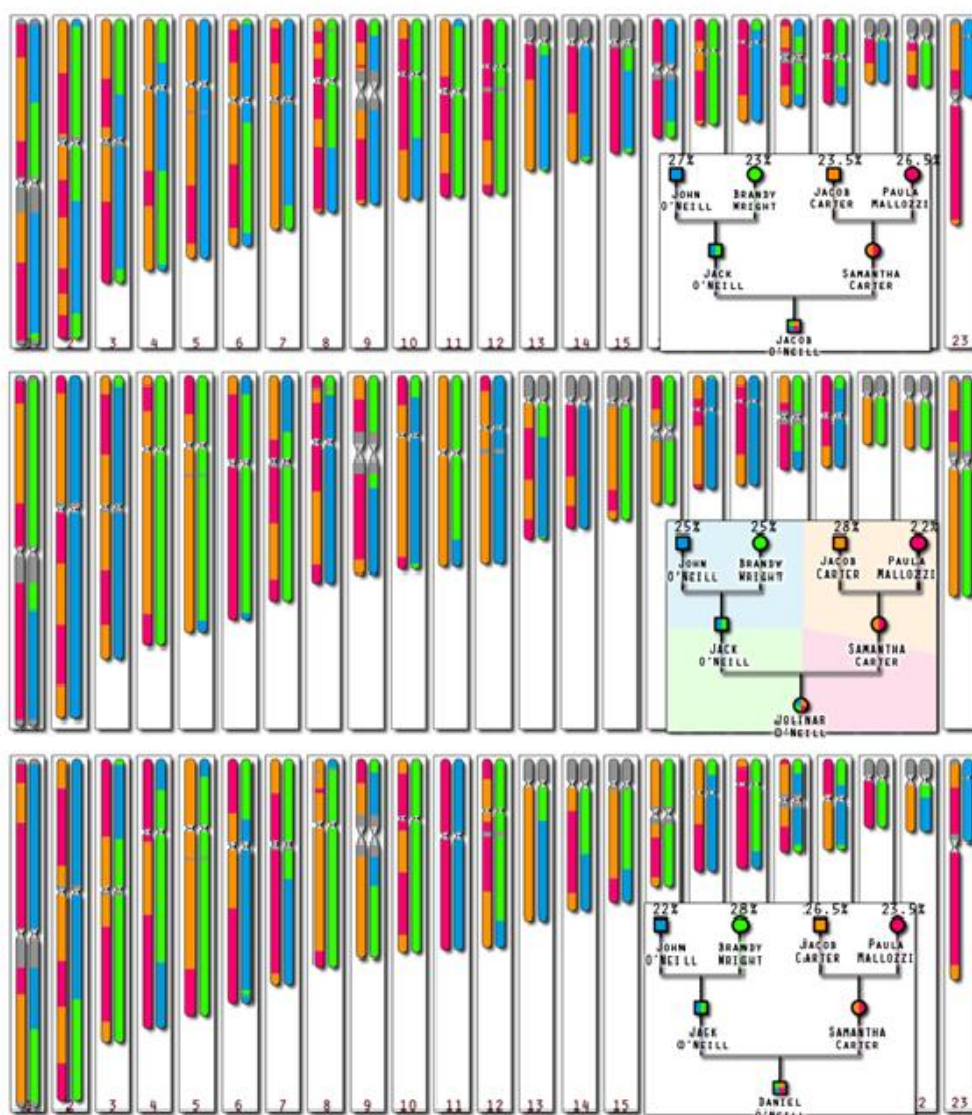
**Figura 3 – Esquema ilustrativo dos fenómenos envolvidos na hereditariedade genética, nomeadamente a gametogénese e a fecundação. As estruturas a cor-de-rosa representam a gametogénese feminina e as azuis a gametogénese masculina. 1 – célula-mãe; 2 – replicação do DNA; 3 e 4 – recombinação ou crossing over; 5 – primeira divisão celular; 6 – segunda divisão celular; 7 – fecundação; 8 – ovo ou zigoto. Fonte: CONE (2010).**

Sob o ponto de vista da variabilidade genética, um dos fenómenos que ocorrem durante o processo de meiose que assume especial relevância é a ocorrência de recombinação (ou *crossing over*). Este fenómeno é o responsável pelo facto de irmãos, mesmo sendo do mesmo sexo (desde que não sejam gémeos homozigóticos) apresentarem diferentes características, visto que cada gâmeta possui uma combinação única de metade dos genes de cada progenitor. Em teoria, como cada indivíduo herda um conjunto de 23 cromossomas do pai e outro tanto da mãe (50% de cada progenitor), a percentagem relativamente aos avós seria de 25%, porém, na realidade não é isso que se observa. Na Figura 4 constam três esquemas onde estão representados os mapas genéticos dos cromossomas de três irmãos e a percentagem do material genético que cada um herdou dos pais e avós. A percentagem de material genético herdado de cada um dos avós, no

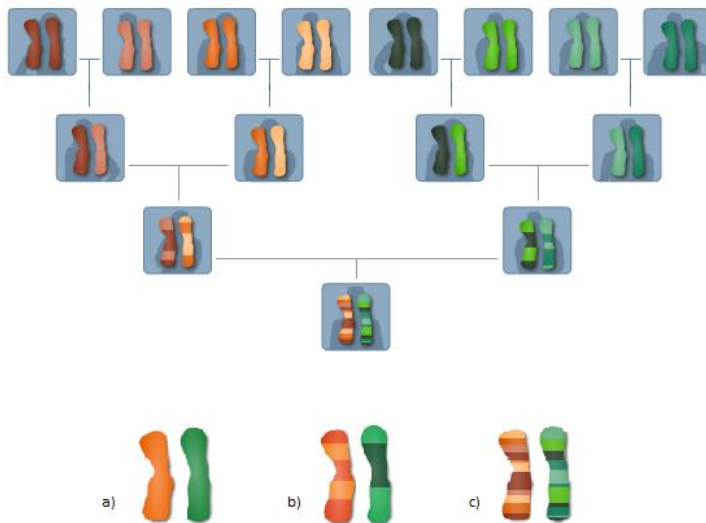
exemplo em análise, varia entre os 22 e os 28%. Esta variação resulta da recombinação que ocorre na meiose (como foi já referido) e na junção aleatória dos gametas na fecundação.

À medida que se observam relações genéticas entre gerações mais distantes, as porções cromossômicas comuns são cada vez menores, como mostra a Figura 5.

O estudo de marcadores autossômicos é utilizado, em contexto forense, no âmbito da identificação e testes de paternidade, mas também no campo da previsão de ancestralidade e características fenotípicas, como será abordado adiante.



**Figura 4 – Mapas genéticos de três irmãos Jacob, Jolinard e Daniel O'Neil. As diferentes cores representam as porções herdadas dos avós. Nas árvores genealógicas adjacentes a cada um dos esquemas encontram-se, para além da descrição da representação das cores do esquema, as proporções de herança do material genético. Fonte: CONE (2010).**

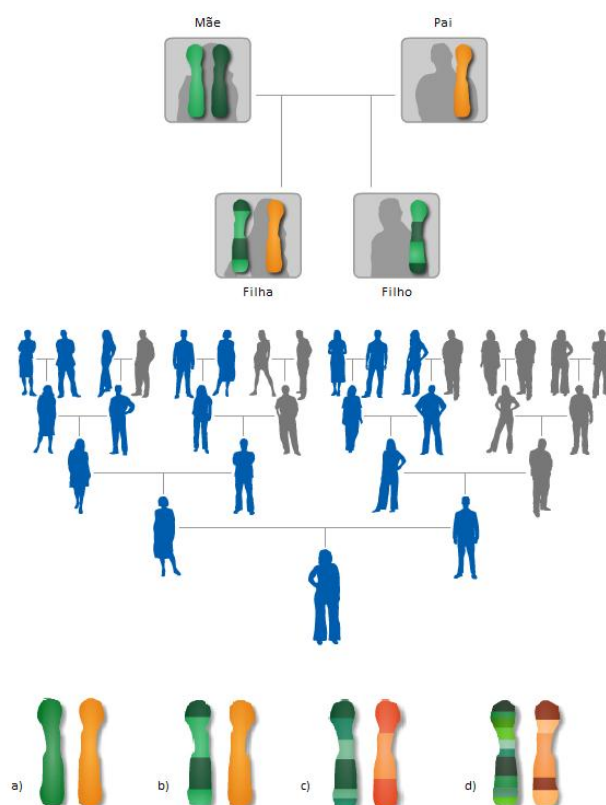


**Figura 5 – Esquema ilustrativo do modo de herança de um par de autossomas. Em cima – Árvore genealógica com representação de quatro gerações, sendo que se encontram representadas com diferentes cores as porções cromossômicas que são transmitidas a partir da primeira geração. Em baixo – esquema representativo de um par de cromossomos no qual as diferentes cores representam a contribuição genética dos antepassados, retrocedendo uma (a), duas (b) e três (c) gerações. Fonte: (THE UNIVERSITY OF UTAH, 2011).**

### *1.1.2. MARCADORES DO CROMOSSOMA X*

O cromossoma X é um dos cromossomas que determinam o sexo do indivíduo. Como foi anteriormente referido, os indivíduos do sexo feminino apresentam um par de cromossomas X, logo, nas mulheres o cromossoma X sofre recombinação e é transmitido à descendência tal como os autossomas (Figura 6, em cima). Nos homens, uma vez que o cromossoma X apresenta apenas uma cópia e não emparelha com o Y (à exceção de uma pequena porção), não ocorre recombinação, sendo que o cromossoma X é transmitido intacto para as descendentes do sexo feminino (visto que os descendentes do sexo masculino herdam o cromossoma Y do pai e não o X). Como consequência, o cromossoma X apresenta um modo de herança diferente dos autossomas (Figura 6, ao centro), sendo que as mulheres herdam um cromossoma X da mãe e um do pai, cromossomas esses que recombina antes de serem transmitidos a todos os seus descendentes, enquanto os homens recebem apenas um cromossoma X da mãe que é posteriormente transmitido exclusivamente aos descendentes do sexo feminino sem sofrer recombinação. Pode assim concluir-se que, uma vez que o grau de recombinação ao qual está sujeito o cromossoma X é menor que os autossomas, o nível de segmentação observado é, consequentemente, menor, com é possível observar na Figura 6, em baixo.





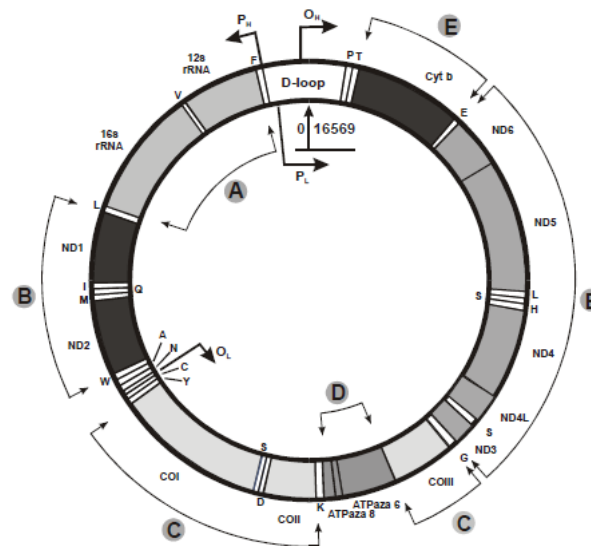
**Figura 6 – Esquemas ilustrativos do modo de transmissão do cromossoma X. Em cima – árvore genealógica representativa do modo de transmissão do cromossoma X onde estão representadas apenas duas gerações; é possível verificar que no caso da mãe o cromossoma X sofre recombinação antes de ser transmitido à descendência ao contrário do que se observa no pai. Ao centro – árvore genealógica onde está representado o modo de transmissão do cromossoma X ao longo de cinco gerações. Em baixo – Representação esquemática de um par de cromossomas X na qual as diferentes cores representam a contribuição genética dos antepassados, retrocedendo uma (a), duas (b), três (c) e quatro (d) gerações. Fonte: (THE UNIVERSITY OF UTAH, 2011).**

Em contexto forense existem alguns cenários nos quais a análise de marcadores do cromossoma X é muito útil, especialmente em casos de testes de paternidade. SZIBOR et al. (2003 e 2007) referem alguns exemplos: (1) pretende-se determinar se uma menina é filha de um homem (pai putativo) mas só se pode testar o DNA da putativa avó paterna; neste caso, a probabilidade de correspondência de marcadores do cromossoma X é de 50%. (2) casos de alegado incesto que resultem no nascimento de uma menina; uma vez que o cromossoma X que o pai transmite aos descendentes é o mesmo, a filha e a neta/putativa filha apresentam um cromossoma X idêntico. (3) casos em que uma menina pode ser filha de um de dois homens, sendo que os putativos pais são pai e filho; neste caso, a análise de marcadores de cromossoma X é muito útil visto que esse cromossoma é diferente nos dois homens pois é sempre herdada da mãe. (4) pretende-se determinar se um homem é pai de uma menina mas apenas o DNA de uma filha do pai putativo está disponível para análise; visto que o cromossoma X é transmitido de forma integral de pai para filha, caso as duas meninas sejam irmãs (filhas do mesmo pai) herdaram o mesmo cromossoma.

### 1.1.3. MARCADORES MITOCONDRIAIS

Relativamente ao DNA mitocondrial (mtDNA), este apresenta-se sob a forma de molécula circular de DNA em cadeia dupla com 16.569 pares de bases (Figura 7) e localiza-se na matriz mitocondrial, estando, por vezes, ligado à membrana interna da mitocôndria (VIDEIRA, 1999). O genoma mitocondrial inclui 37 regiões codificantes (que codificam duas moléculas de rRNA, 22 moléculas de tRNA e 13 proteínas envolvidas no processo de fosforilação oxidativa), sendo que todos os transcritos e seus produtos de tradução permanecem no organelo, não se observando, portanto, exportação de RNA ou proteínas para o citoplasma (VAN OVEN et al., 2009, VIDEIRA, 1999).

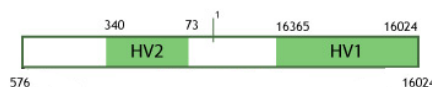
**Figura 7 - Representação esquemática do DNA mitocondrial humano. A - genes de RNA ribossomal; B - genes do complexo I (NADH desidrogenase); C - genes do complexo IV (citocromo C oxidase); E - genes do complexo III (ubiquinol-citocromo C oxidorreductase). Fonte: WITAS et al. (2004).**



As posições dos nucleótidos do mtDNA humano encontram-se numeradas de 1 a 16.569, por convenção, de acordo com a Sequência de Referência de Cambridge revista (rCRS) (ANDERSON et al., 1981, ANDREWS et al., 1999). A essa sequência de referência são posteriormente anotadas as diferenças detetadas (polimorfismos do tipo SNP ou mutações). O nível de variação não é constante ao longo da molécula de mtDNA, sendo que a taxa de variação na região de controlo (região não-codificante situada entre as bases 16.024 e 576 – Figura 8) é cerca de 10 vezes superior à observada na região codificante (região situada entre as bases 577 e 16.023) (VAN OVEN et al., 2009). Esta diferença é explicada pelo facto de a ocorrência de mutações em regiões codificantes ter como consequência alterações ao nível dos produtos de transcrição, e, por conseguinte, ao nível do funcionamento de processos vitais ao funcionamento celular. Relativamente à região de

controle, apesar de desempenhar um importante papel ao nível da replicação do mtDNA, observa-se, como foi anteriormente referido, uma maior taxa de variação, especialmente ao nível das regiões hipervariáveis (HV1 e HV2). No entanto, por não se tratar de uma região codificante, a maioria dessas mutações não tem consequências práticas ao nível do funcionamento celular, sendo permitida a sua acumulação nesta região.

As regiões hipervariáveis presentes na região de controlo (D-loop) caracterizam-se pela presença de um elevado número de polimorfismos do tipo SNP que se associam em estruturas designadas haplótipos (conjunto de marcadores genéticos intimamente ligados com localização próxima uns dos outros que tendem a ser herdados e transmitidos em conjunto/bloco, não sofrendo recombinação frequentemente). As regiões hipervariáveis (HV1 e HV2) (Figura 8) são, como o próprio nome indica, as regiões do mtDNA que apresentam maior variabilidade, no entanto, também na região codificante é possível encontrar SNPs que podem ser úteis na diferenciação de amostras que são semelhantes ao nível dos marcadores da região de controlo (NATIONAL FORENSIC SCIENCE TECHNOLOGY CENTER).

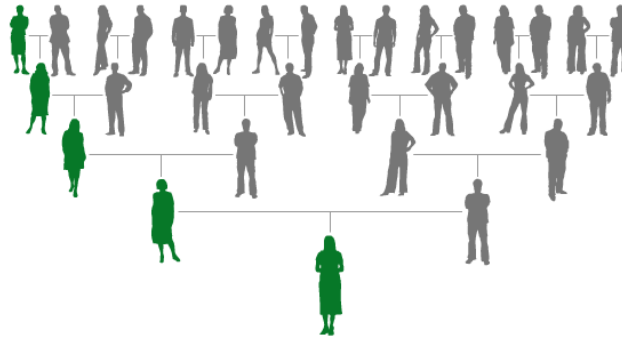


**Figura 8 – Esquema representativo da região de controlo (D-loop) onde se encontram discriminadas as regiões hipervariáveis (1 e 2). Adaptado de NATIONAL FORENSIC SCIENCE TECHNOLOGY CENTER.**

Relativamente ao modo de herança, visto que as mitocôndrias que formam o zigoto são provenientes do gameta feminino, o mtDNA de cada indivíduo é proveniente da mãe (salvo raras exceções nas quais se provou a herança de mtDNA paterno (SCHWARTZ et al., 2002) ou casos em que se observou a presença de mais do que um tipo de mtDNA num indivíduo devido à ocorrência de mutações (LI et al., 2010), fenómeno que toma o nome de heteroplasmia).

O mtDNA não sofre (em regra) recombinação, pelo que a análise deste tipo de material genético permite aceder diretamente à linhagem materna do indivíduo (Figura 9).



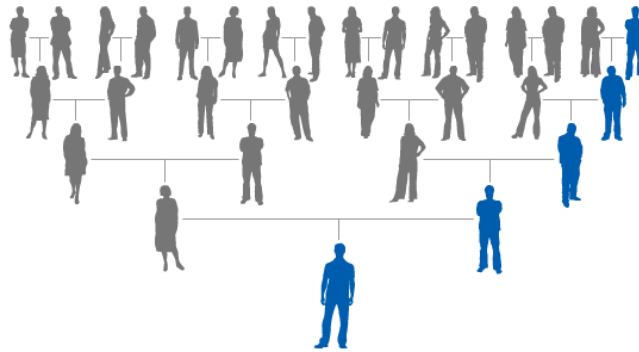


**Figura 9 – Árvore genealógica onde se encontra representado o modo de transmissão do mtDNA. Fonte: THE UNIVERSITY OF UTAH (2011).**

Uma importante vantagem do mtDNA é o facto de existir um elevado número (que pode variar entre as centenas e os milhares) de cópias por célula, dependendo do tipo de célula, enquanto os autossomas apresentam apenas duas cópias por célula e a porção não recombinante do cromossoma Y (NRY) apresenta apenas uma cópia por célula. Este facto torna a análise de mtDNA extremamente vantajosa no campo forense nos casos de análise de material *post-mortem*, em que os processos de degradação e quantidade limitada de material em análise limitam ou impossibilitam a análise de marcadores autossomais ou da NRY. Assim, a análise de mtDNA assume um papel preponderante na identificação de restos mortais humanos realizada a partir de ossos, cabelo ou dentes, sendo aplicada na identificação de vítimas de desastres em massa ou vítimas de guerra através da comparação com eventuais familiares maternos vivos (KAYSER, 2007).

#### **1.1.4. MARCADORES DO CROMOSSOMA Y**

O cromossoma Y é exclusivo dos elementos masculinos da população, sendo transmitido de pai para filho (Figura 10) Ao contrário dos restantes cromossomas, é constitutivamente haplóide, sendo que apenas dois segmentos (regiões pseudossomais) recombinam com o cromossoma X, o que representam cerca de 5 % do comprimento total do cromossoma (JOBLING et al., 2003). No entanto, quando se fala do cromossoma Y no contexto de estudos moleculares a referência diz geralmente respeito à porção do cromossoma que não sofre recombinação relativamente à qual existem, segundo JOBLING et al. (2003), várias possíveis designações: Y não recombinante (NRY), porção não recombinante do Y (NRPY) ou Y específico masculino (MSY).



**Figura 10 - Árvore genealógica onde se encontra representado o modo de transmissão do cromossoma Y. Fonte THE UNIVERSITY OF UTAH (2011).**

O estudo de marcadores do cromossoma Y tem, de acordo com (JOBILING et al., 2003), várias aplicações no campo forense, nomeadamente na análise de amostras de casos de violação e em casos de determinação de parentesco (de meninos) nos quais não se disponha de amostra de DNA do pai mas esteja disponível material genético de um familiar do sexo masculino de linhagem paterna (como por exemplo o avô paterno ou um irmão do pai).

## 1.2. POLIMORFISMOS DE DNA QUE SÃO USADOS COMO MARCADORES GENÉTICOS

A miríade de pequenas e grandes diferenças na sequência nucleotídica entre os indivíduos são conhecidas como polimorfismos de DNA. Um polimorfismo de DNA corresponde, então, à existência simultânea de mais do que uma forma sequencial num determinado locus numa determinada população.

Na área forense, o primeiro método de estudo da variabilidade genética com a finalidade de realizar a identificação de indivíduos foi o RFLP (*restriction fragment length polymorphism* – polimorfismo de comprimento de fragmento de restrição). Neste método é realizada a tipagem de regiões do DNA nas quais se observam polimorfismos do tipo VNTR (*variable number of tandem repeats* – número variável de repetições em tandem). Este método foi posteriormente substituído por métodos com amplificação de loci de DNA por PCR (*polymerase chain reaction* – reação de polimerização de cadeia) seguida da tipagem de marcadores genéticos de STR (*short tandem repeat* – curtas repetições em

tandem). O novo método veio permitir a análise de amostras degradadas e de quantidade limitada, uma vez que a PCR permite aumentar o número de cópias das várias regiões de interesse (marcadores) e o comprimento dos fragmentos STR é menor do que os VNTR (BUDOWLE et al., 2008). Para além disso, o número de marcadores em estudo e as suas características permitem um maior poder de discriminação, como será abordado de seguida.

### *1.2.1. SHORT TANDEM REPEATS (STRs)*

Os *Short Tandem Repeats* (STRs) são também designados microssatélites e consistem em sequências em que se observa a repetição de unidades de um a seis nucleótidos (mono-, di-, tri-, tetra-, penta-, ou hexanucleótidos). Alguns autores (INTERNATIONAL HUMAN GENOME SEQUENCING CONSORTIUM, 2001) defendem que os microssatélites podem apresentar entre 1 a 13 nucleótidos nas unidades repetitivas. Quando o número de nucleótidos presentes nas unidades repetitivas é superior passa a estar-se na presença de um minissatélite (13 a 500 nucleótidos segundo o INTERNATIONAL HUMAN GENOME SEQUENCING CONSORTIUM (2001)) ou, em última instância, um satélite de DNA quando o número de nucleótidos ultrapassa os 500 (BIRREN et al., 1997, ELLEGREN, 2004).

Segundo um estudo do INTERNATIONAL HUMAN GENOME SEQUENCING CONSORTIUM (2001) com base nos dados obtidos através da primeira sequenciação do genoma humano, os STR representam cerca de 3% do genoma humano (Tabela 3), sendo que as repetições dinucleotídicas são as mais frequentes (0,5%). Dentro das repetições dinucleotídicas, as mais frequentes são a AC (50%) e AT (35%), sendo a GC a menos frequente (0,1%). As repetições trinucleotídicas mais frequentes são a AAT e a AAC (33% e 21%, respetivamente) enquanto que as unidades repetitivas menos observadas, dentro deste tipo, são a ACT (1,4%) e a ACG (0,1%). Este estudo conclui que, de uma forma geral, os microssatélites com dois nucleótidos são mais frequentes do que os de três nucleótidos (Tabela 3).

**Tabela 3 – Teor de STRs do genoma humano. Como é observável, os STRs mais comuns no genoma humano são as repetições de dinucleótidos. Informação adaptada de: INTERNATIONAL HUMAN GENOME SEQUENCING CONSORTIUM (2001).**

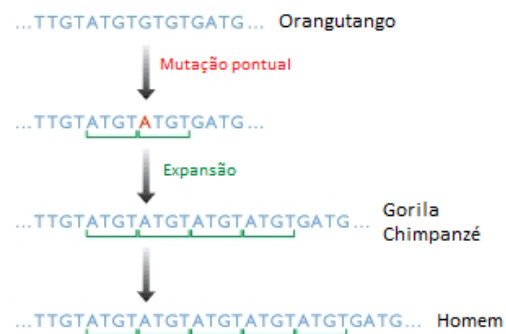
<b>Comprimento da unidade de repetição (bases)</b>	<b>Média de bases por Mb</b>	<b>Número médio de elementos STR por Mb</b>
1	1.660	36,7
2	5.046	43,1
3	1.013	11,8
4	3.383	32,5
5	2.686	17,6
6	1.376	15,2
7	906	8,4
8	1.139	11,1
9	900	8,6
10	1.576	8,6
11	770	8,7

Os microssatélites ocorrem com maior frequência em regiões não-codificantes, quer nas regiões intergênicas, quer nos intrões. Assume-se que nestas regiões os STRs evoluem de forma neutra por não estarem sujeitos a pressões seletivas, sendo que os microssatélites usados como marcadores são normalmente deste tipo. Nas regiões codificantes a atuação da seleção natural contra mutações que originem *frameshift* dificulta o aparecimento de microssatélites é exceção das repetições de trinucleótidos (ELLENGREN, 2004). Este tipo de repetições não altera a forma como a sequência é posteriormente traduzida mas tem muitas vezes consequências ao nível do funcionamento da proteína codificada. A doença de Huntington, a atrofia muscular bolbo espinal (doença de Kennedy) e a ataxia espinocerebelar do tipo I são exemplos de doenças neurodegenerativas provocadas pela expansão (aumento do número de unidades repetitivas) de trinucleótidos “CAG” (BATES et al., 1994).

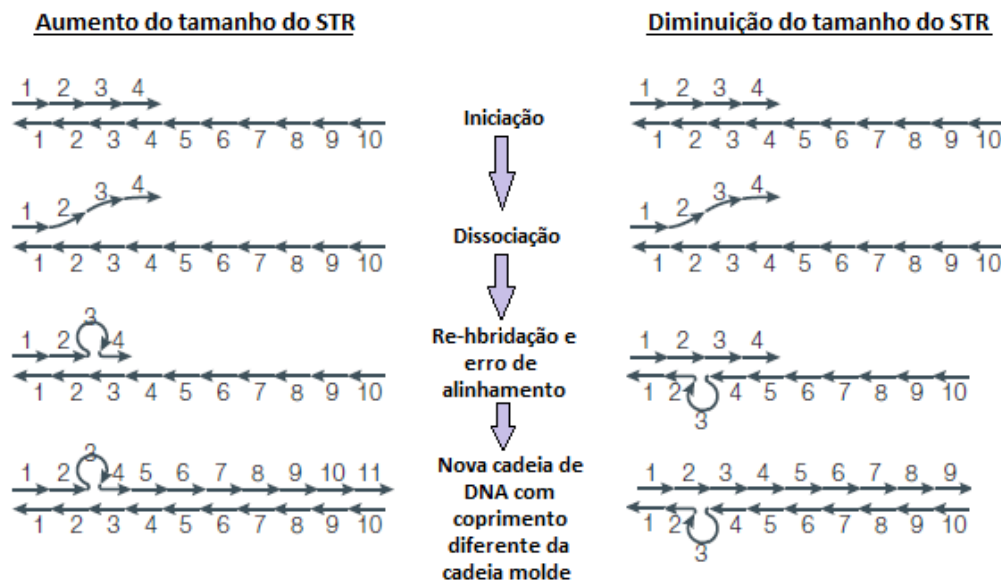
#### 1.2.1.1. A ORIGEM DE UM STR

A origem de um microssatélite pode ser consequência de um de dois tipos de mutações: mutações pontuais que originam duas ou três unidades repetitivas ou inserções de segmentos de DNA que são repetições da sequência adjacente. Um exemplo do primeiro mecanismo referido foi documentado por MESSIER et al. (1996) através da comparação da sequência de nucleótidos de um pseudogene de globulina de espécies de primatas e humanos (Figura 11). A sequência 5'-ATGTGTGT3' ocorre numa posição relevante em espécies de primatas com alguma distância filogenética dos humanos. Relativamente aos humanos e espécies filogeneticamente mais próximas, como o gorila e o chimpanzé, verificou-se a ocorrência de uma mutação pontual que modificou esta sequência para 5'-ATGTATGT3', com subsequente expansão da unidade de repetição 5'-ATGT3' através do processo de derrapagem de replicação (BROWN, 2002).

**Figura 11** – Esquema representativo da origem de um microssatélite documentada através da comparação da sequência de nucleótidos observada em diferentes espécies. Adaptado de: BROWN (2002).



A derrapagem de replicação consiste na ocorrência de um erro ao nível do alinhamento aquando da re-hibridização após a dissociação transitória do DNA em processo de replicação, formando-se uma ansa de DNA, como mostra a Figura 12. Deste fenómeno pode resultar o aumento ou diminuição do número de unidades repetitivas que compõem o microssatélite consoante a ansa se formar na cadeia nascente ou na cadeia molde de DNA, respetivamente. O sistema de reparação que atua no sentido de corrigir erros que ocorrem no processo de replicação de DNA repara a maioria destas mutações, sendo que apenas uma pequena percentagem destas variações persiste (ELLEGRÉN, 2004).



**Figura 12 - Esquema ilustrativo do processo de derrapagem de replicação.** Após a iniciação do processo de replicação de DNA há a possibilidade de ocorrer dissociação entre as cadeias; se, quando da re-hibridização das cadeias, ocorrer um erro de alinhamento numa região onde se observe a ocorrência de um microssatélite que dê origem a uma ansa ou *loop* na cadeia nascente, o processo resultará no aumento do comprimento do STR; por outro lado, se o *loop* se formar na cadeia molde, o resultado será a diminuição do comprimento do STR. Figura adaptada de ELLEGREN (2004).

À medida que um STR se expande ao longo do tempo, a ocorrência de uma mutação pontual ao longo da sua extensão é uma óbvia possibilidade e, quando ocorre, verifica-se uma irregularidade ao nível da sequência repetitiva que se traduz na alteração da estrutura ou composição do polimorfismo e na diminuição da taxa de mutação daquele alelo específico (BRINKMANN et al., 1998). Assim, a evolução dos microssatélites é um processo dinâmico, sendo que tanto o aumento ou a diminuição do número de repetições que os constituem como a variação da composição são característicos desse mesmo processo de evolução (ELLEGREN, 2004).

#### 1.2.1.2. CARACTERÍSTICAS DOS MARCADORES DE STR

Os STRs apresentam uma frequência elevada e distribuição aleatória no genoma. (BIRREN et al., 1997). Para além disso, estes polimorfismos apresentam uma taxa de mutação que é diretamente proporcional ao número de unidades de repetição, o que é

previsível visto que quanto maior o número de unidades repetitivas maior a possibilidade de ocorrência de derrapagem de replicação. A elevada taxa de mutação que caracteriza os microsatélites leva, ao longo do tempo, a que se tornem altamente polimórficos dentro da população e, conseqüentemente, altamente informativos. Um STR é classificado como polimórfico se o número de unidades de repetição variar dentro da população com uma incidência superior a 1% (BIRREN et al., 1997, ELLEGREN, 2004, PANZER et al., 1995).

Contudo, o facto de apresentarem uma elevada taxa de mutação nem sempre constitui uma vantagem, uma vez que marcadores particularmente suscetíveis a mutações podem dificultar o mapeamento de traços por estudos de desequilíbrio de ligação (LD) (BIRREN et al., 1997). Outro facto relevante diz respeito ao relativamente reduzido comprimento que estes polimorfismos apresentam, o que torna possível a amplificação de fragmentos de DNA por PCR que contenham toda a extensão do STR e a sua posterior análise por eletroforese (em gel ou capilar), o que permite detetar diferenças de tamanho dos produtos de amplificação gerados, provenientes de diferentes fontes de DNA genómico (BIRREN et al., 1997, DALE et al., 2002, PANZER et al., 1995).

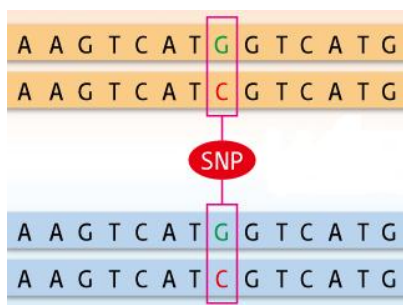
Assim, as vantagens principais dos STRs são o seu elevado poder informativo como marcadores genéticos, a conveniência da sua deteção por métodos baseados em PCR (sendo que é necessária uma quantidade relativamente pequena de DNA) e a possibilidade de transferir informação sobre os marcadores com base apenas em sequência de primers de PCR (BIRREN et al., 1997). Estas vantagens fazem dos STRs marcadores de eleição na identificação forense. A utilização de painéis de STRs (sob a forma de kits multiplex que amplificam vários loci em simultâneo) permite a obtenção de perfis que apresentam probabilidades de correspondência aleatória inferiores a  $10^{-23}$  (JOBILING et al., 2004). O desenvolvimento desses painéis em conjunto com os avanços técnicos, nomeadamente ao nível das tecnologias de sequenciação (em particular a eletroforese capilar), tornou possível o desenvolvimento de bases de dados de DNA nacionais que constituem importantes ferramentas na investigação forense (PHILLIPS, 2008).

Relativamente ao uso destes marcadores genéticos em estudos populacionais há que referir que o facto de se observarem taxas de mutação bastante heterogéneas entre diversos loci (e até entre alelos do mesmo locus (BRINKMANN et al., 1998)) torna difícil a tarefa de traduzir estimativas de distância genética em escalas temporais absolutas (ELLEGREN, 2004).

### 1.2.2. SINGLE NUCLEOTIDE POLYMORPHISMS (SNPs)

Cerca de 85 a 90% das variações ao nível da sequência de DNA corresponde a diferenças num único nucleótido (substituições, deleções ou inserções) e são designadas *Single Nucleotide Polimorphisms* (SNPs) (BIRREN et al., 1997, BUDOWLE et al., 2008, BUDOWLE et al., 2009, COLLINS et al., 1998). Alguns autores defendem que os SNPs correspondem apenas a polimorfismos de substituição (Figura 13), sendo que as inserções e deleções constituem um grupo separado designado “indels” (GONZALEZ et al., 2007, MILLS et al., 2006, MULLANEY et al., 2010). MULLANEY et al. (2010) define os “indels” como inserções ou deleções de 1 a 10.000 pares de bases de comprimento. Por seu turno, (GONZALEZ et al., 2007) definem “indel” como “uma mutação que resulta na deleção e inserção co-localizadas e no ganho ou perda de nucleótidos” adicionando ainda o termo “microindel” que, segundo os autores, corresponde a “um indel que resulta na perda ou ganho de 1 a 50 nucleótidos”. Assim, regra geral, as substituições e deleções são estudadas separadamente das substituições, considerando-se apenas as substituições nas referências aos SNPs.

A maioria dos SNPs apresenta apenas dois alelos, e por isso os SNPs são muitas vezes referidos como marcadores bialélicos. Com apenas dois alelos, o nível máximo de heterozigotia para cada marcador é de 50%, o que faz com que os SNPs sejam menos informativos do que os STRs, que tipicamente apresentam vários alelos e níveis de heterozigotia a rondar os 70% (BIRREN et al., 1997, BUDOWLE et al., 2009).

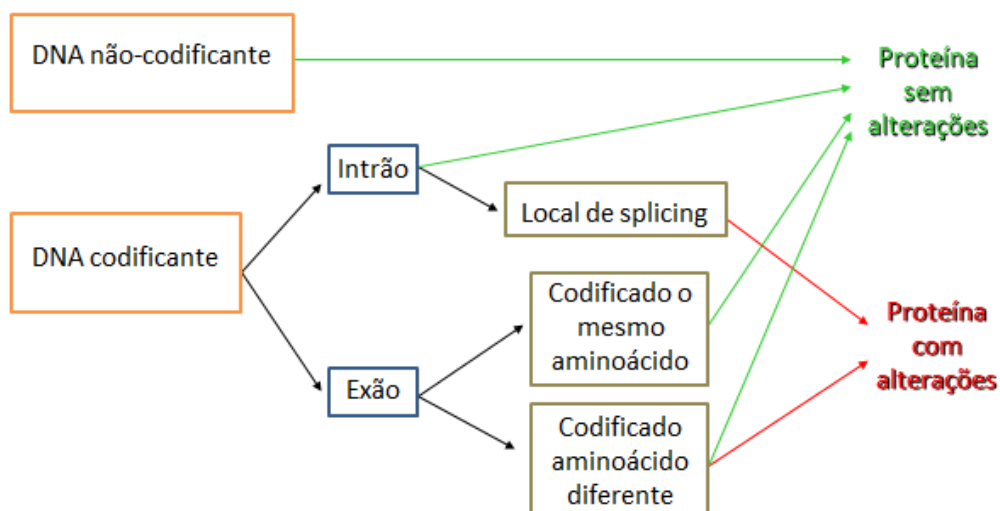


**Figura 13 – Esquema representativo da forma mais comum de SNP: substituição de um nucleótido. Adaptado de: [http://www.rikenresearch.riken.jp/images/figures/hi\\_3675.jpg](http://www.rikenresearch.riken.jp/images/figures/hi_3675.jpg) consultado a 29-06-2010.**

As consequências da alteração de uma única base na sequência de DNA (SNP) são variáveis (Figura 14), dependendo do local onde esta ocorre. Em primeiro lugar, a alteração pode ocorrer numa região não codificante; de facto, nos seres eucariotas, esta é a ocorrência mais comum (DALE et al., 2002, TISHKOFF et al., 2004). Este fenómeno deve-se



ao facto de a maior parte do DNA do genoma das células eucariotas corresponder a sequências não codificantes. Por outro lado, um polimorfismo que ocorra numa região não codificante apresenta uma menor probabilidade de causar efeitos negativos no funcionamento celular. Em segundo lugar, se a alteração ocorrer numa sequência codificante, podemos considerar dois cenários distintos: (1) a alteração ocorre num intrão – excetuando os casos em que a alteração ocorre no local de splicing, não há alteração no produto do gene; (2) a alteração ocorre num exão – mesmo neste caso pode não ocorrer alteração na sequência da proteína codificada pelo gene em causa. Uma vez que o código genético é redundante, ou seja, um aminoácido pode ser codificado por tripletos diferentes, quando a substituição de uma base origina um triplete diferente, mas que codifica o mesmo aminoácido, as consequências ao nível do funcionamento da célula serão nulas, embora possa haver algum efeito na tradução do mRNA (como por exemplo, a leucina pode ser codificada pelo UUA e CUA), sendo que esta alteração é designada de substituição sinónima ou mutação silenciosa. Outras alterações podem também ser silenciosas, embora não sinónimas. A alteração do UUA por GUA iria substituir uma leucina por uma valina na proteína, o que pode representar consequências mínimas na estrutura e funcionamento da proteína (dependendo de quão crítica é a localização específica em causa), uma vez que a leucina e a valina são aminoácidos similares. Outras alterações, como o UUA pelo UCA (serina), têm uma probabilidade maior de comportar consequências mais graves ao nível estrutural e funcional da proteína, embora esta e outras alterações ainda mais radicais possam ser toleradas em algumas posições (DALE et al., 2002).



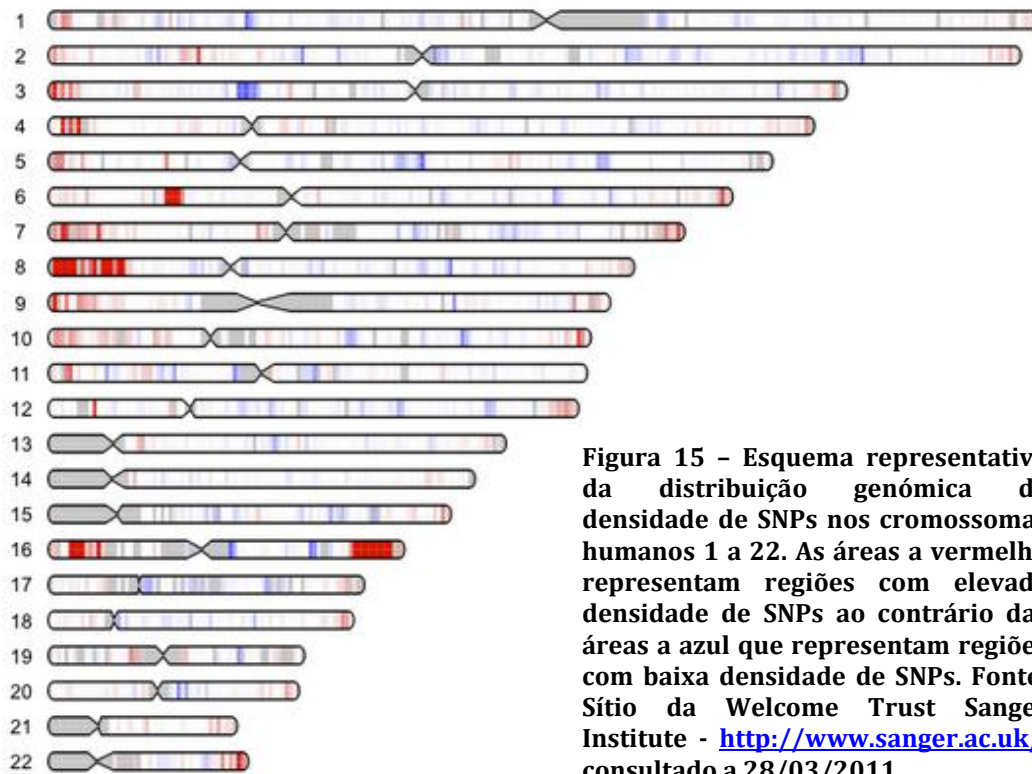
**Figura 14 – Esquema ilustrativo das possíveis consequências da ocorrência de SNPs em diversos locais do genoma.**

Como foi referido anteriormente, os SNPs podem ou não comprometer a estrutura e função de proteínas. Pontualmente, a ocorrência deste tipo de variação genética pode levar a uma alteração na função da proteína – por exemplo, uma enzima pode ser capaz de atuar sobre um substrato diferente. Um exemplo de região codificante que apresenta um elevado nível de variabilidade é o sistema HLA (antigénio leucocitário humano) (THE 1000 GENOMES PROJECT CONSORTIUM, 2010). Esta região representa o Complexo Major de Histocompatibilidade (MHC) nos seres humanos e localiza-se no braço curto do cromossoma 6. As proteínas codificadas pelos vários genes deste complexo exercem funções ao nível da apresentação de antígenos às células C (sistema imunitário) e moléculas intervenientes no processo inflamatório (GOLDSBY et al., 2004). Segundo dados de Julho de 2011 da base de dados IMGT/HLA, foram já identificados mais de 6.800 formas alélicas deste complexo (EUROPEAN BIOINFORMATIC INSTITUTE, 2011). A elevada variabilidade observada nesta região faz com que a sua análise seja útil não só em termos clínicos, mas também ao nível do estudo da história evolutiva do MHC humano e da variação genética entre populações (BUHLER et al., 2011).

Por outro lado existem também regiões do genoma implicadas no processamento de RNA ou na regulação da transcrição e desenvolvimento que se encontram altamente conservadas (inclusivamente ao nível não só da espécie, mas de todos os mamíferos), ou seja, o nível de variação observado é extremamente baixo (BEJERANO et al., 2004).

Assim, é de prever que a distribuição dos SNPs no genoma não seja homogénea. De facto, observam-se regiões nas quais a densidade deste tipo de polimorfismos é maior (por exemplo, junto aos telómeros, ou seja, perto das extremidades dos cromossomas) como pode observar-se na Figura 15.

O facto de os SNPs ocorrerem com grande frequência no genoma humano, faz com que constituam uma fonte valiosa de marcadores genéticos para testes de identidade, mapeamento de características simples ou complexas, estudos de associação genótipo-fenótipo e reconstrução da evolução humana.

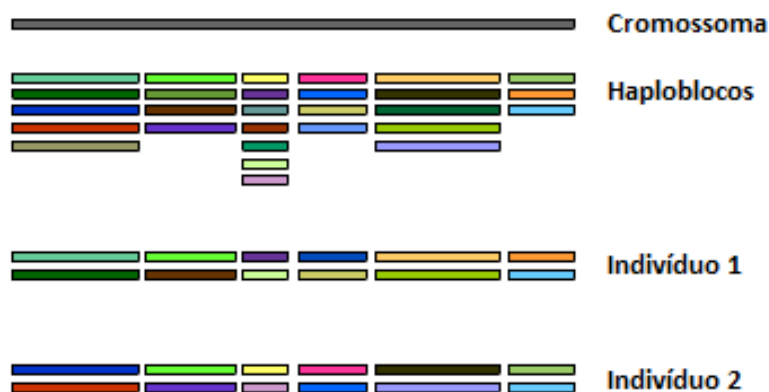


**Figura 15 – Esquema representativo da distribuição genómica da densidade de SNPs nos cromossomas humanos 1 a 22. As áreas a vermelho representam regiões com elevada densidade de SNPs ao contrário das áreas a azul que representam regiões com baixa densidade de SNPs. Fonte: Sítio da Wellcome Trust Sanger Institute - <http://www.sanger.ac.uk/>, consultado a 28/03/2011.**

#### 1.2.2.1. DESEQUILÍBRIO DE LIGAÇÃO E HAPLÓTIPOS

Por se localizarem fisicamente próximos uns dos outros, alguns marcadores são tendencialmente transmitidos em bloco, o que significa que à presença de um alelo num determinado locus está associada a presença de um alelo específico noutro locus próximo. Os haplótipos são, então, conjuntos de marcadores intimamente ligados que são geralmente herdados em conjunto, ou, por outras palavras, são combinações de estados alélicos de um restrito número de loci próximos uns dos outros. Assim, os loci pertencentes a um determinado haplótipo encontram-se em desequilíbrio de ligação (LD).

Outro conceito importante a referir é o conceito de “haplogrupo” (Figura 16). Um haplogrupo é então a combinação de vários haplótipos e corresponde a um tipo de estrutura que assume especial relevância no estudo do DNA mitocondrial e do cromossoma Y, como será abordado adiante.



**Figura 16 – Esquema ilustrativo da organização do material genético em haploblocos em cromossomas de dois indivíduos diferentes. Adaptado de: <http://www.dnai.org/d/index.html> consultado a 07-04-2011.**

O LD verifica-se, como foi já referido, quando a presença de um alelo num determinado locus está associada à presença de um outro alelo específico num locus próximo, o que faz com a que a probabilidade de ocorrência dos dois alelos considerados não seja independente (THE INTERNATIONAL HAPMAP CONSORTIUM, 2005). Assim sendo, pode dizer-se que o LD constitui uma medida indireta da distância entre loci (JORDE et al., 2001). Ao longo do tempo, o LD sofre uma diminuição que é uma consequência da ocorrência de recombinação, o que faz com que o padrão de LD reflita várias gerações de recombinação uma vez que é potencialmente afetado por todas as recombinações que ocorreram ao longo da história (JORDE et al., 2001). Outros fatores que têm influência nos padrões de LD são a distância física interlocus, a distância ao centrómero, a presença de *hotspots* de recombinação, a localização no cromossoma, a população em estudo, a conversão génica e o tipo de marcadores usados (JORDE et al., 2001)

Há ainda que ter em consideração outros fatores que, no contexto da história evolutiva humana, afetam os padrões de LD – mutações, fluxo de genes, deriva génica, crescimento da população e subdivisão da população (JORDE et al., 2001). TISHKOFF et al. (2004) refere que os níveis de LD são mais elevados nas populações com origem mais recente, pela seguinte ordem: Nativo-Americanos > Eurasiáticos > Africanos, sendo que a variação dos níveis de LD também se faz notar dentro das próprias populações. Os autores referem também que esta variação é consistente com a história demográfica humana,

nomeadamente com o efeito de “gargalo de garrafa” que ocorreu aquando das duas migrações intercontinentais, efeito esse que se refletiu, então, nas quantidades crescentemente elevadas de LD verificadas nas populações mais recentes.

#### 1.2.2.2. CLASSIFICAÇÃO DOS SNPs EM CONTEXTO FORENSE

Segundo (BUDOWLE et al., 2008), os SNPs com aplicação na área forense podem ser classificados em quatro categorias: os SNPs de identificação, os SNPs informativos de fenótipo, SNPs informativos de ancestralidade e os SNPs informativos de linhagem.

Os SNPs de identificação são usados para individualizar, ou seja, funcionam como os atuais marcadores STR de identificação. Os SNPs informativos de fenótipo são usados para estabelecer a probabilidade de um indivíduo apresentar uma determinada característica fenotípica, como a cor do cabelo, pele ou olhos que forneçam pistas acerca do dador da amostra. Em relação aos SNPs informativos de ancestralidade, estes são usados no estabelecimento da ancestralidade biogeográfica de um indivíduo no sentido de inferir indiretamente algumas características fenotípicas que forneçam pistas acerca do suspeito. Por fim, os SNPs informativos de linhagem correspondem a conjuntos de SNPs intimamente associados que funcionam como marcadores de haplótipos para identificar pessoas desaparecidas através de análise de parentesco, sendo também analisados no contexto da avaliação da ancestralidade biogeográfica dos indivíduos.

Os diversos tipos de SNPs (em particular os SNPs informativos de fenótipo, os SNPs informativos de ancestralidade e os SNPs informativos de linhagem) são abordados com maior pormenor adiante.

#### 1.2.3. SNPs vs STRs

Uma vez que a maioria dos SNPs são bialélicos, estes são menos informativos para efeitos de identificação do que a análise de STRs. (BIRREN et al., 1997, BUDOWLE et al., 2008, BUDOWLE et al., 2009). Estima-se que a quantidade de SNPs necessária para obter o

mesmo poder de discriminação que o fornecido pelos sistemas multiplex de STRs seria de 50-100 marcadores. Para realizar a tipagem de tal bateria de marcadores, seria necessária uma quantidade de amostra de DNA muito maior do que a necessária para os sistemas multiplex de STRs (BUDOWLE et al., 2008). Para além disso, o facto de a maioria dos SNPs serem bialélicos levantaria problemas ao nível da interpretação em situações que envolvessem mistura de amostras (ou contaminação). Apesar de alguns protocolos de análise de SNPs permitirem aferir concentrações relativas dos dois alelos, ajudando a reconhecer amostras misturadas, FRUDAKIS (2008) afirma que é pouco provável que os SNPs venham a substituir os STRs em casos de amostras misturadas quando o objetivo é a identificação. O autor acrescenta que nos casos de violação é também possível separar as frações de DNA feminino e masculino por lise diferencial de células. Para além disso, tendo em conta que as bases de dados de DNA para uso forense têm por base os marcadores STR, a possibilidade de que os SNPs venham a substituir os STRs como marcadores forenses principais (em particular no campo da identificação) a curto ou médio prazo é muito remota, de acordo com GILL et al. (2004).

Tendo em conta o que foi anteriormente referido pode dizer-se que os STRs são mais vantajosos do que os SNPs no campo da identificação, no entanto, segundo FRUDAKIS (2008), no campo da previsão de ancestralidade biogeográfica e fenótipos os SNPs apresentam mais vantagens. Segundo o autor, os SNPs encontram-se na base da maior parte da variação funcional, constituindo a maior porção da variação genética entre humanos. Uma vez que os SNPs são frequentes em regiões codificantes ou próximo das mesmas, e uma vez que os genes influenciam de forma direta as características fenotípicas, os SNPs são então relevantes na determinação de fenótipos. Em contraste, os STRs apresentam uma relativamente baixa prevalência e apenas alguns destes polimorfismos encontram-se associados a efeitos funcionais, como, por exemplo, na doença de Huntington e síndrome do X frágil. Para além disso, a baixa taxa de mutação associada aos SNPs contribui para efeitos de subestrutura da população, o que os torna úteis na determinação da variação entre populações e consequentemente na previsão de ancestralidade biogeográfica (BUDOWLE et al., 2008, FRUDAKIS, 2008). Um estudo publicado por LAO et al. (2006) refere que um painel de 10 SNPs autossómicos selecionado pelos autores permite obter resultados comparáveis aos obtidos por ROSERBERG et al. (2002) recorrendo a 377 STRs, o que prova o facto de os SNPs (desde que convenientemente selecionados) serem mais informativos relativamente à ancestralidade biogeográfica.

Estes polimorfismos são também mais numerosos, mais estáveis e potencialmente mais fáceis de classificar do que os STRs (COLLINS et al., 1998).

Na Tabela 4 encontram-se resumidas as principais características que distinguem os STRs e os SNPs.

**Tabela 4 – Principais características que distinguem os STRs e os SNPs no âmbito da análise forense de DNA. Adaptado de BUTLER (2009).**

Características	STRs	SNPs
Ocorrência no genoma humano	~1 a cada 15 kb	~1 a cada 1 kb
Conteúdo informativo geral	Elevado	Baixo: cerca de 20 a 30% relativamente aos STRs
Taxa de mutação	~1 em 1000	~1 em 100.000.000
Tipo de marcador	Marcadores de repetições di-, tri-, tetra- e pentanucleótidos com vários alelos	Normalmente marcadores bialélicos com seis possibilidades: A/G, C/T, A/T, C/G, T/G A/C
Número de alelos por marcador	Normalmente 5-20	Normalmente 2 (existem alguns SNPs trialélicos)
Capacidade de previsão de origem biogeográfica	Limitada	Sim
Informação fenotípica	Não	Sim
Principais vantagens para aplicação forense	Grande número de alelos permitem taxas de sucesso mais elevadas de deteção e interpretação de misturas	Produtos de amplificação podem ser pequenos, potenciando a obtenção de melhores resultados em amostras degradadas

### 1.3. MARCADORES GENÉTICOS QUE REFLETEM A DIVERSIDADE INTERPOPULACIONAL

Sendo o *Homo sapiens sapiens* uma espécie relativamente recente (cerca de 200.000 anos (JURMAIN et al., 2008)), observa-se uma diversidade genética muito limitada relativamente a outras espécies (CONRAD et al., 2007, FRUDAKIS, 2008). No entanto, várias são as características que assumem diferentes aspetos em diferentes populações,

sejam características físicas (tom de pele, cor dos olhos e do cabelo, características faciais, forma do cabelo, massa corporal, altura,) ou características fisiológicas (resposta a fármacos, suscetibilidade e resistência a doenças) (WU et al., 2011). Essas características são traços complexos, sendo, por isso, produto da interação entre fatores genéticos e ambientais. É, então, de prever, que existam diferenças ao nível genético entre indivíduos oriundos de diferentes populações.

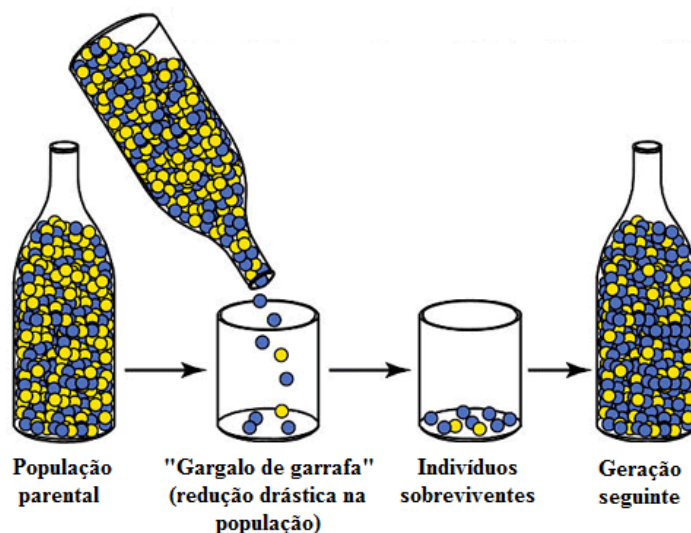
Os processos (forças evolutivas) que estão na origem da variabilidade genética interpopulacional são a mutação, a deriva genética, a mistura e a seleção natural (FRUDAKIS, 2008, POLLARD et al., 2006).

A mutação corresponde a uma alteração na sequência, número ou posição de ácidos nucleicos que compõe a cadeia de DNA. As mutações ocorrem ao longo de praticamente todo o genoma e constituem a principal fonte de variabilidade genética, sendo, por isso, importantes como uma força evolutiva a longo prazo, como quando se fazem comparações entre espécies. No entanto, as restantes três forças evolutivas são mais importantes quando se pretende avaliar frequências de alelos já existentes, como é o caso (FRUDAKIS, 2008).

Relativamente à deriva genética, trata-se de um fenómeno decorrente da transmissão aleatória do material genético de uma geração para a seguinte. Por outras palavras, os gametas produzidos pelos indivíduos em idade de procriar recebem aleatoriamente uma das duas cópias existentes de cada locus (estes dados são referentes aos autossomas); visto que apenas uma fração dos gametas produzidos se vai unir aos gametas do sexo oposto dando origem a um novo indivíduo, a geração seguinte representa apenas uma pequena amostra de todas as combinações possíveis de todos os gametas produzidos na geração anterior; este processo aleatório de amostragem pode originar flutuações ao nível das frequências alélicas observadas, podendo, ao longo das gerações, resultar na extinção ou fixação de alelos na população. A magnitude do efeito da deriva na variação genética é inversamente proporcional ao tamanho da população em idade de procriar. A perda de alguns alelos e a fixação de outros na população como produto da atuação da deriva genética leva à diferenciação genética entre populações que se tornaram reprodutivamente isoladas. Os níveis de diferenciação são determinados pelo tempo decorrido desde a separação e o tamanho das populações filhas. Há, no entanto, a ter em conta o facto de que, nas populações reais, raramente se observa um isolamento completo (FRUDAKIS, 2008).



Ainda relativamente à deriva genética, existem dois efeitos que se fazem sentir ao nível da diversidade e frequências alélicas que importa referir – efeito “gargalo de garrafa” e efeito fundador. Quando o número de indivíduos de uma população sofre uma drástica diminuição (devido a um desastre natural, predação ou redução de habitat), ocorre uma perda importante da diversidade genética do conjunto de genes (*gene pool*) original – este é o efeito “gargalo de garrafa” (Figura 17). Relativamente ao efeito fundador, este verifica-se quando um grupo constituído por um número reduzido de indivíduos abandona a população original e estabelece uma nova colónia. Neste caso, este grupo representa uma amostra da diversidade genética da população original, sendo que essa amostra pode inclusivamente não ser aleatória.

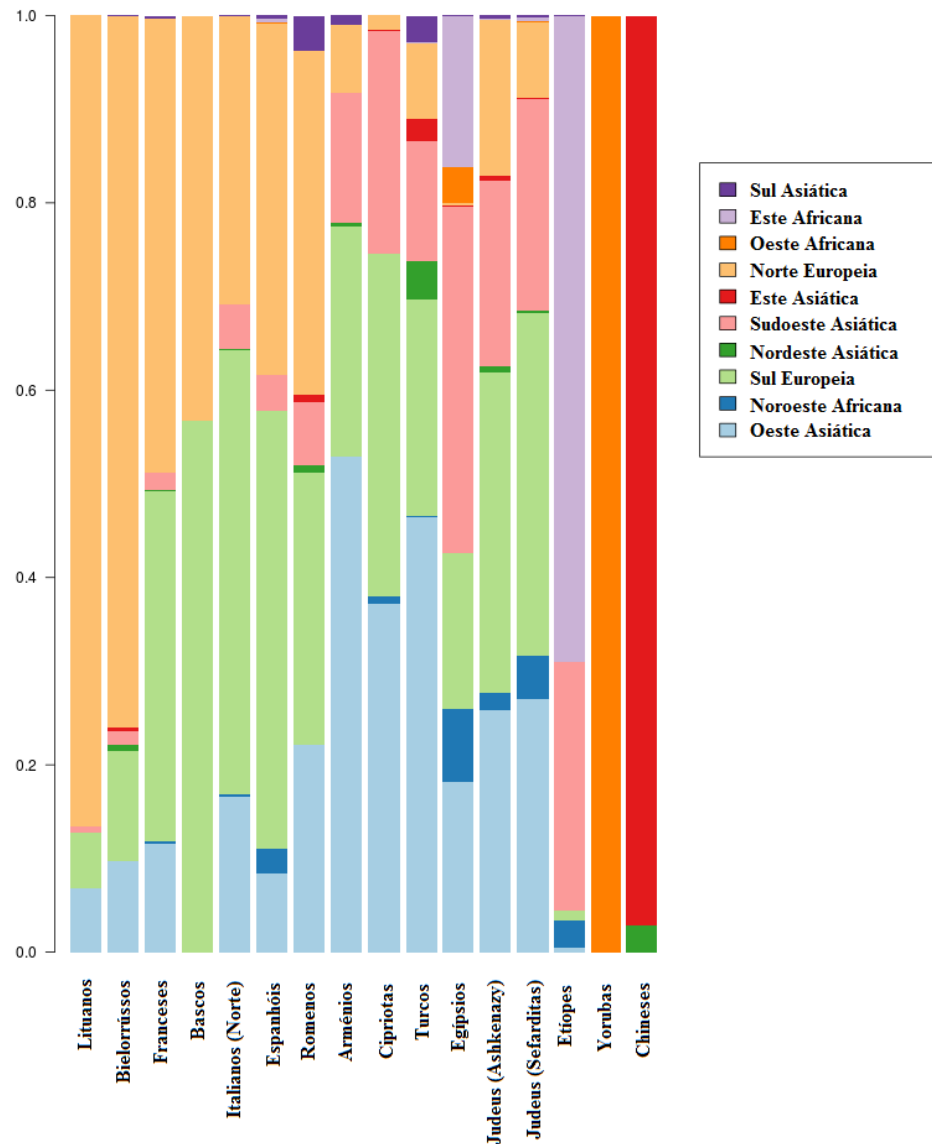


**Figura 17 – Representação esquemática do efeito “gargalo de garrafa”. Uma drástica redução do número de indivíduos que compõem a população devido a um desastre natural, predação ou redução de habitat provoca uma drástica redução na diversidade genética da mesma.**

Figura adaptada de <http://www.biologycorner.com/APbiology/evolution/ch18.html> consultado a 18-10-2011.

A mistura consiste no processo de fluxo ou migração génica de alelos entre populações que ocorre quando indivíduos de populações reprodutivamente isoladas se reproduzem. Este processo contraria os efeitos da deriva genética e da seleção diferencial, tornando as populações mais semelhantes. Relativamente às populações onde ocorre mistura observa-se um fenómeno – estratificação de mistura (Figura 18) – que corresponde à variação na ancestralidade individual existente nos membros da população, ou seja, as proporções de ancestralidade genómica (nível médio de mistura calculado com

base em informações obtidas através do estudo de AIMS) das populações parentais variam de indivíduo para indivíduo (FRUDAKIS, 2008).



**Figura 18** – Gráfico de barras ilustrativo das proporções de mistura de populações ancestrais (representadas pelas diferentes cores) observadas em indivíduos de diferentes países. Gráfico adaptado de uma publicação do blog “Dodecad Ancestry Project” de 23 de Junho de 2011 (<http://dodecad.blogspot.com/2011/06/results-for-dod738-to-dod747-posted.html>) consultado a 18-10-2011.

A seleção natural é o processo através do qual as frequências alélicas são alteradas em função da adaptação diferencial (sobrevivência e reprodução) dos indivíduos de

diferente genótipo. Este processo demarca-se da deriva genética e mistura por afetar apenas a região próxima ao gene que se encontra sob seleção, enquanto os outros dois processos apresentam repercussões ao longo de todo o genoma de forma aleatória. A seleção direcional ou seleção positiva favorece um único fenótipo, o que faz com que as frequências alélicas mudem continuamente nesse sentido, ou seja, a frequência alélica do alelo vantajoso vai aumentar. Um exemplo da ocorrência deste fenômeno é a elevada frequência dos fenótipos de tons claros de olhos e cabelo observada na população Europeia relativamente às outras populações. Este fenômeno observa-se devido à seleção positiva destes fenótipos por serem considerados sexualmente mais atraentes do que os fenótipos de olhos e cabelo escuro o que conferiu uma vantagem adaptativa aos indivíduos com olhos e cabelo claro (FROST, 2006).

Em sentido contrário atua a seleção balanceadora, que, por atuar em favor da heterozigotia, faz com que as frequências alélicas se mantenham em equilíbrio por longos períodos de tempo, tornando mais baixos do que esperados os níveis de diferenciação genética e, assim sendo, favorecendo o aumento da variabilidade genética e fenotípica nas populações. Um exemplo de região codificante que se encontra sob a ação deste tipo de seleção é o MHC. Como foi anteriormente referido, esta região apresenta uma enorme variabilidade. Uma vez que as funções das proteínas codificadas pelos genes incluídos neste complexo desempenham funções ao nível do sistema imunitário, procedendo à apresentação de antígenos (que é um processo essencial na deteção de células infetadas e cancerígenas) a variabilidade observada ao nível dessas proteínas garante uma resposta adequada à grande variabilidade de agentes patogénicos e outras ameaças com que o organismo se pode deparar (ANDRÉS et al., 2010).

Em suma, as quatro forças evolutivas descritas apresentam diferentes características (Tabela 5), principalmente no que diz respeito: (1) à localização das alterações no genoma – as mutações podem ocorrer em qualquer lugar do genoma, embora, existam locais mais propícios à sua ocorrência do que outros; relativamente à deriva genética e mistura, são fenômenos que envolvem todo o genoma de forma aleatória, ao contrário da seleção natural cujo efeito se faz sentir ao nível das regiões codificantes pois são estas que têm influência direta nas características fenotípicas; (2) ao efeito diferencial interpopulacional – relativamente a este assunto, o fator mais importante será a flutuação de frequências alélicas em diferentes populações; assim, a deriva genética e a seleção direcional apresentam um efeito favorável relativamente à diferenciação interpopulacional, enquanto a mistura e a seleção balanceadora levam a uma maior homogeneidade

interpopulacional. Relativamente à mutação, esta força evolutiva apresenta um efeito significativo a longo prazo, apresentando maior interesse na comparação entre espécies do que entre populações da mesma espécie.

**Tabela 5 – Resumo das principais características das forças evolutivas relativamente aos efeitos das mesmas ao nível da diferenciação interpopulacional e da extensão de envolvimento do genoma.**

Forças Evolutivas		Efeito diferencial interpopulacional		Localização das alterações no genoma	
		Aumenta	Diminui	Ubíquas	Localizadas
Mutação		*		X	
Deriva Genética		X		X	
Mistura			X	X	
Seleção Natural	Direcional	X			X
	Balanceadora		X		X

**\* O efeito significativo das mutações faz-se sentir a longo prazo, pelo que esta força evolutiva tem maior interesse na comparação entre espécies; o efeito diferencial interpopulacional relaciona-se diretamente com a variação das frequências alélicas que são significativamente influenciadas pelas outras três forças evolutivas (Deriva Genética, Mistura e Seleção Natural).**

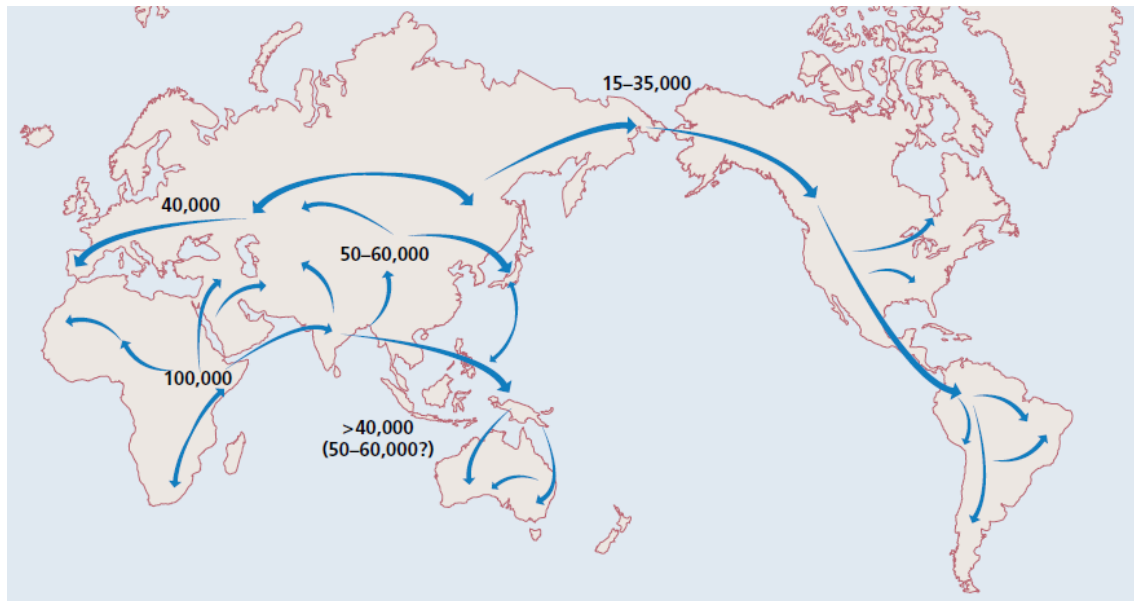
As diferentes forças evolutivas atuaram sobre o genoma humano ao longo da evolução da espécie, moldando a aparência e a fisiologia dos indivíduos de cada população no sentido da adaptação aos diferentes ambientes.

### **1.3.2. ORIGEM E EVOLUÇÃO DA ESPÉCIE HUMANA**

Existem dois modelos principais que explicam a origem e evolução dos seres humanos modernos. Um deles é o modelo de evolução multirregional. Este modelo

defende que as populações de *Homo erectus* que abandonaram África há cerca de 1 milhão de anos atrás para ocupar vários locais do planeta continuaram a sofrer o processo de mudança evolutiva, dando eventualmente origem aos seres humanos modernos em todas as populações do Velho Mundo, tendo ocorrido, durante todo o processo, um extenso fluxo genético entre as diversas populações (LEWIN, 1997). Por outras palavras, este modelo também designado de modelo “Candelabro” defende que os descendentes do *Homo erectus* que migraram a partir de África para os outros continentes formaram populações separadas – os *Homo erectus* Europeus tornaram-se *Homo sapiens* Europeus, os *Homo erectus* Asiáticos tornaram-se *Homo sapiens* Asiáticos e os *Homo erectus* que permaneceram em África tornaram-se *Homo sapiens* Africanos. Cada uma dessas populações evoluiu independentemente devido ao isolamento geográfico, ocorrendo ocasionalmente cruzamentos entre populações, processo este que assegurou uma troca constante de genes entre as regiões. Deste processo resultam populações separadas de humanos modernos, cada uma com características físicas distintas, como é de facto observado nas populações atuais. Em suma, este modelo é caracterizado pela continuidade regional de características anatómicas e defende que os seres humanos modernos apresentam origens variadas, sendo que as populações apresentam raízes genéticas profundas (STANDFORD UNIVERSITY).

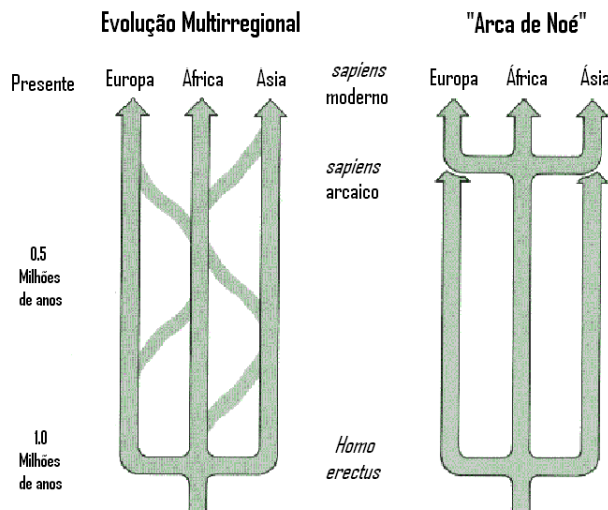
O segundo modelo é designado de modelo “Saída de África” ou modelo “Arca de Noé” e defende que a origem do homem moderno correspondeu a um evento evolutivo discreto com localização definida numa única população africana de *Homo erectus*. Sendo nossos antepassados diretos, estes humanos primitivos possuíam muitas semelhanças connosco, em particular no que toca a características esqueléticas, genéticas e comportamentais (JURMAIN et al., 2008). Os descendentes dessa população de humanos modernos saíram posteriormente de África, há cerca de 100.000 anos ((CAVALLI-SFORZA et al., 1994) citado por (TISHKOFF et al., 2004)), espalharam-se pelo Velho Mundo, substituindo completamente os *Homo erectus* que povoavam os diferentes continentes (LEWIN, 1997). Os padrões de variação alélica indicam que a expansão para novos territórios aconteceu nos seguintes períodos (Figura 19): Austrália há cerca de 60.000 anos, Europa há 40.000 anos, Este da Ásia há 35.000 anos, Continente Americano há 30.000 anos e Pacífico há 3.000 anos ((CAVALLI-SFORZA et al., 1994) citado por (TISHKOFF et al., 2004)).



**Figura 19 – Mapa representativo da migração do *Homo sapiens* moderno onde se encontram representadas as datas de chegada nas respectivas regiões de acordo com dados genéticos. Fonte: CAVALLI-SFORZA et al. (2003).**

Assim, este modelo defende a origem dos seres humanos modernos numa única população de *Homo erectus* Africanos seguida de extensa migração, pelo que as raízes genéticas serão superficiais uma vez que a divisão entre as diversas populações de humanos modernos terá ocorrido mais recentemente do que defende o modelo de evolução multirregional. Relativamente às características anatómicas, este modelo defende que não existe continuidade regional relativamente às mesmas (STANDFORD UNIVERSITY). Na Figura 20 consta um esquema representativo dos dois modelos previamente descritos e na Tabela 6 consta um resumo das principais características que diferenciam os mesmos modelos.

Alguns antropologistas sugerem uma versão menos rígida deste modelo na qual em vez de uma substituição total das populações arcaicas, tenha ocorrido cruzamento entre populações de seres humano modernos e arcaicos (LEWIN, 1997).



**Figura 20** – Esquema ilustrativo de dois modelos da origem dos seres humanos modernos – o modelo de evolução multirregional e o modelo “Arca de Noé”. Adaptado do site: <http://www.stanford.edu/~harryg/protected/evolve4.htm> consultado a 10/01/2011.

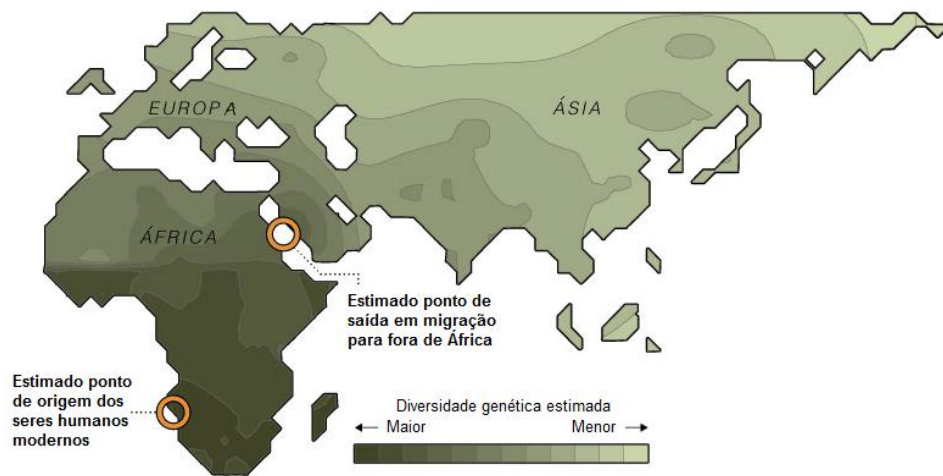
**Tabela 6** – Principais características relativas aos modelos de evolução dos humanos modernos.

Modelo Características	Evolução Multirregional	“Arca de Noé”
Origem dos seres humanos modernos	Múltipla	Única
Raízes genéticas	Profundas	Superficiais
Continuidade regional das características anatómicas	Existente	Não existente

A maioria dos dados anatómicos e arqueológicos apontam para o facto de os seres humanos modernos serem o resultado de um fenómeno evolutivo relativamente recente, facto que vai de encontro ao modelo da “Arca de Noé” (JOHANSON, 2001).

Também os dados obtidos através do estudo de marcadores moleculares apontam no mesmo sentido: os seres humanos são tão semelhantes geneticamente que se observam mais diferenças entre dois chimpanzés oriundos da mesma população do que entre dois humanos da mesma população, bem como se observa uma variação genética ainda maior entre dois chimpanzés de populações diferentes do que entre humanos provenientes de populações diferentes (JOHANSON, 2001, TISHKOFF et al., 2004). De facto, os dados provenientes de estudos moleculares são particularmente importantes no que toca ao seu potencial, se não vejamos: a evolução molecular é marcada pela acumulação de mutações neutras que ocorre a uma taxa constante; sendo a taxa de acumulação dessas mesmas

mutações constante, é possível datar divergências de um tronco comum, ou seja, é possível saber em que período surgiu uma espécie ou se deu a divisão entre populações da mesma espécie (WILSON et al., 1992). O que foi relatado em vários estudos moleculares, envolvendo vários conjuntos de marcadores é que as populações Africanas apresentam um número de alelos específicos maior do que as populações não Africanas, sendo que as segundas são portadoras de apenas uma fração da diversidade genética presente em África (Figura 21) (CANN et al., 1987, JOHANSON, 2001, ROGERS et al., 1992, TISHKOFF et al., 2004, TISHKOFF et al., 2002, WILSON et al., 1992).



**Figura 21 – Mapa representativo da diversidade genética. Adaptado de <http://www.nytimes.com/2009/05/01/science/01eden.html> consultado a 13-09-2011 citando como fontes: Science e Tishkoff et al.**

### *1.3.3. MARCADORES INFORMATIVOS DE ANCESTRALIDADE*

Os Marcadores Informativos de Ancestralidade (AIMs) constituem um grupo de marcadores genéticos com características muito próprias. Trata-se, então, de uma porção inferior a 0,01% do genoma que contém informações relevantes acerca da origem biogeográfica do indivíduo (FRUDAKIS, 2008).

A mais importante característica dos AIMs refere-se à variação observada em termos de frequências alélicas em diferentes populações. Essas diferenças têm por base o processo de dispersão dos seres humanos modernos a partir de África através de dois fatores: o primeiro diz respeito ao efeito fundador verificado sempre que um número restrito de indivíduos deixa a população de origem migrando para um novo espaço



geográfico; a população fundadora é portadora de apenas uma porção da variabilidade genética da população de origem e, em simultâneo, acaba por desenvolver novos perfis genéticos decorrentes de novas mutações (ROYAL et al., 2010). Os factos anteriormente descritos refletem-se nos resultados obtidos através de painéis de marcadores, tanto STRs como SNPs (FRUDAKIS et al., 2003, ROSERBERG et al., 2002). Também a análise da variação interpopulacional em termos de LD revelou que este fenómeno é mais marcado em populações não Africanas o que é um indicador de que estas, por serem mais recentes, apresentam menor influência genética de processos como a mutação e recombinação (TISHKOFF et al., 2004). Por outro lado, fatores como o isolamento geográfico, o acasalamento seletivo (tendência para escolha de parceiro com características fenotípicas semelhantes) e a ocorrência localizada de endogamia com base em fatores socioculturais como a origem étnica e a linguagem contribuíram para o aumento da divergência genética entre populações geograficamente isoladas (CAVALLI-SFORZA, 1997, ROYAL et al., 2010, TISHKOFF et al., 2004).

Uma medida da diversidade genética interpopulacional é o valor de  $F_{ST}$  associado ao marcador. Esta medida estatística de diferenciação genética está diretamente relacionada com a variância da frequência alélica entre populações e, simultaneamente, reflete o grau de semelhança entre indivíduos da mesma população. O  $F_{ST}$  corresponde à correlação entre alelos escolhidos aleatoriamente dentro de uma subpopulação e os alelos da população, o que equivale à proporção da diversidade genética associada a diferenças nas frequências alélicas entre populações (HOLSINGER et al., 2009), sendo definido pela fórmula:

$$F_{ST} = 1 - \frac{H_w}{H_b}$$

onde, segundo HUDSON et al. (1992), “ $H_w$ ” corresponde ao número médio de diferenças entre diferentes sequências amostradas da mesma população, e “ $H_b$ ” corresponde ao número médio de diferenças entre sequências amostradas de duas subpopulações diferentes”.

Esta medida estatística reflete a influência das forças evolutivas na medida em que estas afetam a diferenciação interpopulacional. Neste sentido, o nível de  $F_{ST}$  observado vai diminuir quando se verifica ancestralidade comum recente ou elevados níveis de migração e um elevado tamanho efetivo da população (HOLSINGER et al., 2009, TISHKOFF et al., 2004). Por outras palavras, grandes populações entre as quais ocorre frequentemente

migração tendem a apresentar menor diferenciação (logo, menor valor de  $F_{ST}$ ) do que populações pequenas entre as quais ocorre pouca migração (HOLSINGER et al., 2009). Por outro lado, também a seleção diferencial exerce influência no valor de  $F_{ST}$  observado, sendo que essa influência se faz notar quando se observam valores extremos de  $F_{ST}$  em determinados loci em diferentes estudos de SNPs (TISHKOFF et al., 2004). Por outras palavras, se um determinado alelo de um determinado locus constituir uma vantagem adaptativa numa determinada população, sendo favorecido pela seleção natural, o valor de  $F_{ST}$  nesse mesmo locus será maior do que em loci nos quais as diferenças interpopulacionais são apenas o resultado de deriva genética (HOLSINGER et al., 2009).

Assim, quanto mais elevado o valor de  $F_{ST}$  associado a um marcador, maior a informação que este fornece relativamente à subestrutura da população, sendo que o valor de  $F_{ST}$  pode variar entre 0 (que indica a inexistência de diferenciação) e 1 (que indica total isolamento entre populações) (HOLSINGER et al., 2009, TISHKOFF et al., 2004).

**Tabela 7 – Comparações de valores de  $F_{ST}$  baseada em 369 pares de SNPs\* de populações\*\* ao longo das regiões major. Adaptado de TISHKOFF et al. (2004).**

	<b><math>F_{ST}</math> entre regiões</b>					
	África Europa	África E. Ásia	África América	Europa E. Ásia	Europa América	E. Ásia América
n	48	42	42	56	56	49
Média	0,152	0,228	0,232	0,122	0,156	0,159
Mediana	0,147	0,223	0,229	0,115	0,154	0,161
Mínimo	0,112	0,183	0,171	0,089	0,084	0,086
Máximo	0,202	0,283	0,308	0,182	0,223	0,229

n – número médio de comparações emparelhadas;

\* Os SNPs analisados são maioritariamente não-codificantes mas o estudo inclui algumas substituições silenciosas.

\*\* As amostras incluem indivíduos de um total de 28 populações (6 Africanas, 8 Europeias, 7 Este Asiáticas e 7 Americanas) que se encontram descritas no artigo referido.

Analisando a Tabela 7 é possível observar que, regra geral, o valor médio de  $F_{ST}$ , quando se comparam as populações da região Africana com as das restantes regiões, assume um padrão crescente à medida que são consideradas regiões cada vez mais distantes em termos geográficos.

Outro fator que influencia o valor de  $F_{ST}$  é o tipo de marcador/polimorfismo analisado. De acordo com o trabalho de HOLSINGER et al. (2009), estudos baseados nas

diferenças genéticas entre as regiões geográficas major recorrendo a STRs e SNPs sugerem que a estimativa do valor de  $F_{ST}$  entre as populações amostradas é de 0,05 para os microssatélites de 0,10 para os SNPs. Estes dados reforçam a ideia de que os SNPs são os marcadores mais apropriados ao estudo e, posteriormente, à previsão de ancestralidade biogeográfica do que os STRs defendida por BUDOWLE et al. (2008) e FRUDAKIS (2008), como foi anteriormente referido.

Também o tipo de material genético em análise faz variar os valores de  $F_{ST}$  visto que, segundo TISHKOFF et al. (2004) estimativas deste valor estatístico com base no mtDNA ( $F_{ST} = 0,24 - 0,27$ ) ou no cromossoma Y ( $F_{ST} = 0,23 - 0,64$ ) reportam valores mais elevados do que as estimativas para o DNA autossómico. Os autores defendem que estas diferenças têm por base o reduzido tamanho populacional efetivo do mtDNA e do cromossoma Y comparativamente ao DNA autossómico, facto esse que se reflete no aumento do impacto da deriva genética.

#### 1.3.3.1. HAPLOGRUPOS DA NRY E DO MTDNA

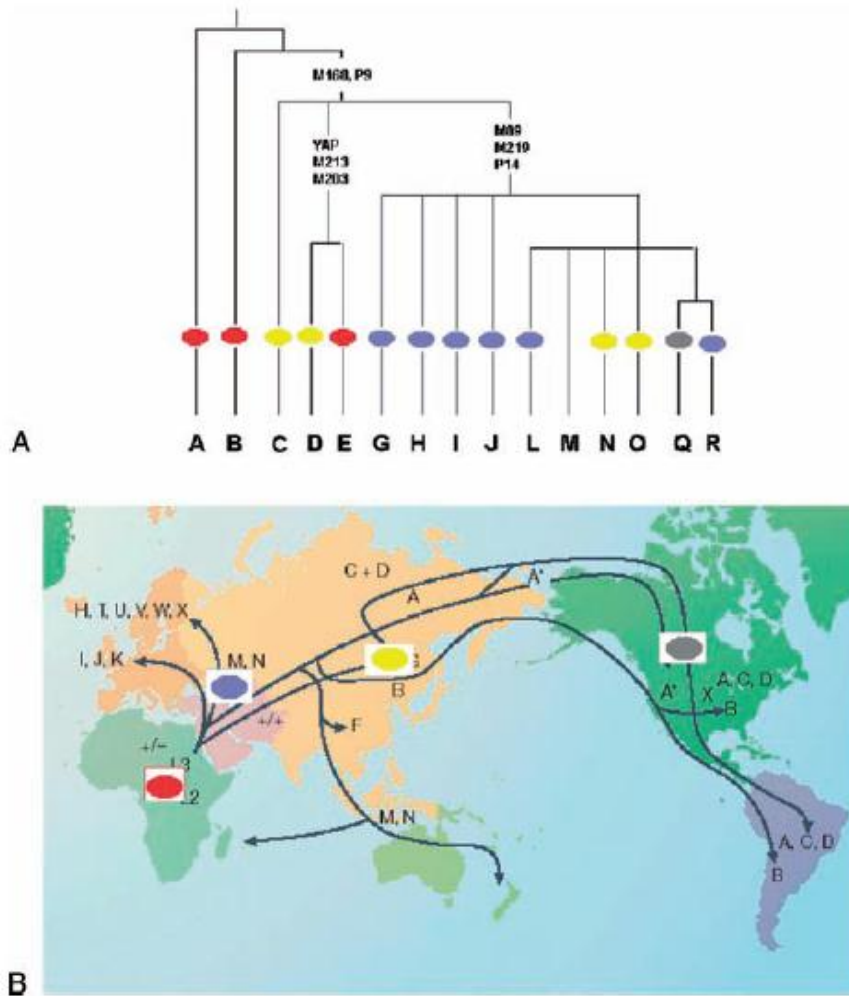
O mtDNA e o cromossoma Y constituem tipos especiais de material genético haplóides não recombinante herdados uniparentalmente, sendo por isso particularmente sensíveis às influências da deriva genética, em especial ao efeito fundador. Observam-se, nestas regiões do genoma humano, estruturas particularmente interessantes – os haplogrupos – que, como foi anteriormente referido, consistem na combinação de vários haplótipos, que, por seu turno, correspondem a combinações de estados alélicos de um restrito número de loci próximos uns dos outros que se encontram em LD. Assim, estes tipos de material genético são ideais para avaliar a diversidade das populações contemporâneas (que pode ser esquematizada sob a forma de árvores filogenéticas - Figura 22 e Figura 23) e inferir relações filogeográficas (JOBILING et al., 2003, UNDERHILL et al., 2001).

Uma das características que distingue o mtDNA e o NRY dos autossomas é o tamanho populacional efetivo ( $N_e$ ). Trata-se de um parâmetro básico em vários modelos de genética populacional que corresponde ao número de indivíduos de uma população ideal (com ausência de seleção e sobreposição de gerações, com igual proporção entre os sexos e tamanho da população constante) cuja variabilidade genética é correspondente à

observada numa população real (FRUDAKIS, 2008). Assim, em relação ao NRY, sendo os homens os portadores do cromossoma Y, e considerando um rácio 1:1 entre homens e mulheres, pode dizer-se que o tamanho populacional efetivo do cromossoma Y corresponde a  $1/4$  do tamanho populacional efetivo de qualquer outro autossoma,  $1/3$  do cromossoma X e é semelhante ao do mtDNA. Assim, observa-se uma menor diversidade ao nível da sequência nucleotídica e uma maior suscetibilidade de ocorrência de deriva genética, que tem como consequência, por um lado, a aceleração da diferenciação entre grupos em diferentes populações, e por outro a possibilidade de ocorrerem rápidas alterações ao nível das frequências de haplótipos ao longo do tempo (FRUDAKIS, 2008, JOBLING et al., 2003).

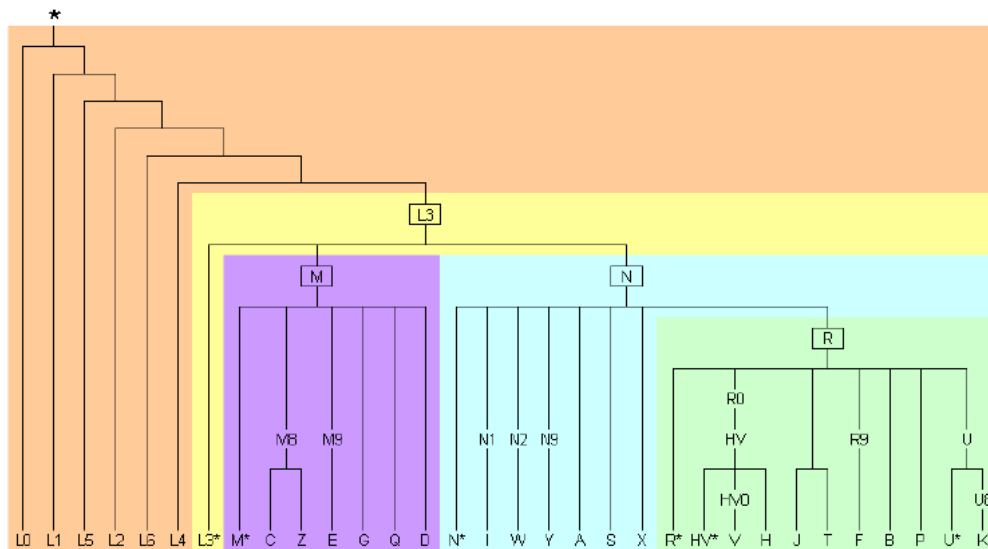
Outro fator que influencia o agrupamento geográfico reside no comportamento dos homens, nomeadamente a questão da patrilocalidade. Este fenómeno é observado em cerca de 70% das sociedades modernas e corresponde ao facto de as mulheres estarem mais frequentemente sujeitas a mobilidade geográfica do que os homens, o que se traduz numa maior diferenciação geográfica dos haplogrupos do cromossoma Y do que do mtDNA (HAMILTON et al., 2005, JOBLING et al., 2003).

Na Figura 22 encontram-se representadas as relações filogenéticas entre os diversos haplogrupos do cromossoma Y e respetivas origens e distribuição atual, sendo que FRUDAKIS (2008) descreve toda a filogenia da NRY em alguns princípios básicos que explicam a maioria da diversidade dos haplogrupos deste cromossoma. Assim, segundo o autor, os Oeste Africanos são portadores dos haplogrupos ancestrais A e B; o haplogrupo E é também característico de populações Africanas, tendo, eventualmente, dado origem a linhagens não-Africanas; os Europeus (incluindo Europa, Médio Oriente e Oeste da Ásia) derivaram dos Africanos, sendo portadores ou tendo adquirido os haplogrupos G, H, I, J, L e R; por seu turno, os Este Asiáticos são portadores ou adquiriram os haplogrupos D, C, N e O; já o haplótipo Q foi adquirido pelos Nativo-Americanos ao passo que os Australianos derivam de fundadores que eram portadores ou adquiriram o haplogrupo M e outros.



**Figura 22 – Esquemas representativos da filogenia do cromossoma Y, correspondendo as cores aos previsíveis locais de origem dos respectivos haplótipos (A) e da distribuição geográfica atual dos haplogrupos (B). Adaptado de FRUDAKIS (2008).**

A acumulação sequencial de mutações ao longo das linhagens maternas está na base da variação observada ao nível da sequência do mtDNA. Assim essa variação pode ser representada sob a forma de árvore que reflete as relações filogenéticas das variantes de mtDNA conhecidas, tal como foi anteriormente referido relativamente à NRY. Os ramos dessa árvore filogenética (Figura 23) são definidos por haplogrupos que tomam designações de “A” a “Z” por ordem de descoberta. Os Haplogrupos L constituem as linhagens mais antigas e encontram-se restritas ao território Africano (o que é consistente com a teoria da “arca de Noé”). Em relação às linhagens observadas fora de África, todas elas têm origem nos macrohaplogrupos M, N e R (sendo o R um subgrupo do N), que, por seu turno tiveram origem no Haplogrupa L3 (VAN OVEN et al., 2009).



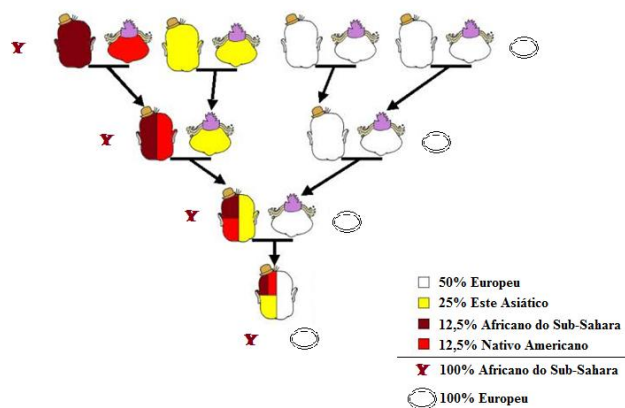
**Figura 23 – Árvore filogenética simplificada onde se encontram os diversos haplogrupos do mtDNA representados pelas respectivas letras; \*) outras linhagens não representadas; Adaptado de VAN OVEN et al. (2009).**

As filogenias da NRY e mtDNA estão em constante atualização visto serem descobertas novas variantes com grande frequência. Assim, o *Y Chromosome Consortium* (YCC) - <http://ycc.biosci.arizona.edu/> constitui um grupo que se dedica ao estudo da variação genética do cromossoma Y, disponibilizando informação regularmente atualizada relativamente a novos polimorfismos do NRY e mudanças na nomenclatura dos haplogrupos e filogenia (CALAFELL et al., 2002). Já a base de dados YHRD (Y Chromosome Haplotype Reference Database - <http://www.yhrd.org/>) armazena informação relativa aos haplótipos do cromossoma Y (WILLUWEIT et al., 2007). Em relação ao mtDNA, a base de dados EMPOP (<http://empop.org/>) corresponde a uma base de dados desenvolvida pelo Instituto de Medicina Legal e Instituto de Matemática da Universidade de Innsbruck e visa a recolha e controlo de qualidade de dados dos haplótipos da região de controlo do mtDNA de todo o mundo (PARSON et al., 2007). Por seu turno, o site <http://www.phylotree.org/> disponibiliza informação regularmente atualizada relativamente aos haplogrupos com base em informação da região de controlo e da região codificante (VAN OVEN et al., 2009).

Na interpretação dos dados genéticos relativos a marcadores da NRY e mtDNA há que ter em atenção o facto de estes dados não serem representativos de toda a

variabilidade genética do indivíduo, particularmente em termos de ancestralidade biogeográfica, uma vez que se trata de material genético herdado uniparentalmente. Assim, por hipótese, um indivíduo do sexo masculino cuja mãe e respetivos progenitores sejam de origem Europeia e o pai de origem Africana, Este Asiática e Nativo-Americana apresentará um perfil biogeográfico relativamente aos marcadores autossomais que reflete as proporções de origem biogeográfica dos seus antepassados. Relativamente aos marcadores do NRY e mtDNA, visto não sofrerem recombinação e serem transmitidos uniparentalmente, permitem apenas aceder a informação relativa à linhagem paterna e materna, respetivamente, como se encontra esquematizado na Figura 24.

Assim, determinar o haplogrupo ao qual um indivíduo pertence não é sinónimo de determinar a sua origem biogeográfica mas sim apenas uma parte da mesma (PHILLIPS et al., 2009) pelo que a interpretação dos resultados obtidos na análise de marcadores de ancestralidade da NRY e mtDNA deve ser muito cuidadosa.



**Figura 24 – Árvore genealógica na qual estão representados os diferentes modos de herança dos marcadores autossómicos (nas figuras da árvore), NRY ("Y") e mtDNA ("○") cujas proporções relativamente ao indivíduo que está na base da árvore se encontram descritas na legenda. Figura adaptada de <http://stevemorse.org/genetealogy/dna.htm> consultado a 18-10-2011**





# CAPÍTULO 2

## 2. METODOLOGIAS DE ANÁLISE DE SNPs

A limitada quantidade de material biológico presente nas amostras forenses, bem como o elevado nível de degradação muitas vezes observado nas mesmas, implica a utilização de métodos e marcadores especialmente sensíveis na análise forense do material genético. Nesse sentido, a análise de SNPs tem vindo a demonstrar muitas potencialidades em especial pelo facto de exigir fragmentos íntegros de DNA com dimensões reduzidas comparativamente aos STRs (HARRISON et al., 2006).

De acordo com SOBRINO et al. (2005), a análise de SNPs implica o recurso a tecnologia e metodologias específicas que devem obedecer a determinados critérios, nomeadamente a elevada sensibilidade, precisão e reprodutibilidade, a capacidade de genotipar vários loci em simultâneo (sistema multiplex) e com elevado nível de rendimento (número de marcadores genotipados em cada ciclo). Os autores defendem que o nível de rendimento necessário pode variar consoante as atividades específicas de cada laboratório: para casos de investigação criminal e investigação de paternidade é suficiente um método de rendimento médio; porém, aquando da implementação de bases de dados de DNA torna-se necessário o recurso a métodos de genotipagem de elevado rendimento que permitam a análise de um grande número de amostras. Outros critérios apontados pelos autores dizem respeito ao tempo de realização da genotipagem, custo (quer dos equipamentos quer das próprias genotipagens) e a flexibilidade da tecnologia (possibilidade de utilização dos equipamentos na realização de outros métodos). KWOK (2001) acrescenta outro ponto importante: o método deve ser o mais automatizado possível e a análise dos dados deve ser simples e precisa. O autor aborda três aspetos da genotipagem de DNA que diferenciam os diferentes métodos que se encontram descritos na Tabela 8 relativamente a seis metodologias de genotipagem com aplicação no campo

forense (*Reverse dot blots*, PCR quantitativa em tempo real, pirosequenciação, minissequenciação, MALDI TOF MS e Microarrays).

Relativamente à discriminação alélica (e tendo em conta que apenas se incluem os métodos referidos) pode ser realizada através de hibridização ou extensão de primer. No primeiro caso, são usadas sondas alelo-específicas que hibridizam com a sequência alvo apenas em caso de total correspondência enquanto que o processo de extensão de primer implica o recurso a primers que hibridizam com a sequência alvo imediatamente antes do local onde ocorre o SNP, identificando-se posteriormente qual a base que foi incorporada.

No que toca ao formato da reação, o autor discrimina entre reações homogêneas e reações com suporte sólido, sendo que as primeiras têm lugar em meio líquido enquanto nas segundas as sondas encontram-se ancoradas a um suporte sólido como uma membrana de nylon (*Reverse dot blots*) uma microesfera (pirosequenciação) ou uma lâmina de vidro ou silicone (*Microarrays*).

O último aspeto focado na tabela são os mecanismos de deteção, sendo que, em relação aos seis métodos em abordagem, foram identificados cinco métodos de deteção diferentes: deteção de fluorescência, FRET (*fluorescence resonance energy transfer* – transferência de energia de ressonância de fluorescência), deteção de luminescência e espectrometria de massa. Os métodos que fazem uso da deteção de fluorescência recorrem à incorporação de nucleótidos com marcação fluorescente que é posteriormente identificada através de eletroforese capilar (minissequenciação) ou, no caso dos *microarrays* de DNA, os produtos de amplificação que se pretende genotipar são marcados com fluorescência, sendo que, quando se dá a hibridização entre estes e as sondas presentes no *microarray* é detetada, posteriormente, a emissão de fluorescência. Nos métodos que recorrem à PCR quantitativa em tempo real a deteção da fluorescência pode ser direta (o fluoróforo emite fluorescência quando integrado na dupla cadeia de DNA – SYBR Green) ou, noutros casos, é baseada na transferência de energia de ressonância de fluorescência (FRET). Este mecanismo corresponde à transferência de energia que se observa entre dois elementos (fluoróforo-quencher) quando estes se encontram próximos no caso do TaqMan®.

Tabela 8 – Tabela onde estão representados as principais metodologias utilizados na genotipagem de SNPs no âmbito forense e respectivas características relativamente à discriminação alélica, formato da reação e mecanismos de detecção. Informação baseada no trabalho de KWOK (2001).

Aspetos técnicos da genotipagem de DNA		Metodologias de genotipagem de SNPs					
		<i>Reverse dot blots</i>	PCR quantitativa em tempo real	Pirossequenciação	Minissequenciação	MALDI TOF MS	<i>Microarrays</i>
Discriminação alélica	Hibridização	X	X				X
	Extensão de primer			X	X	X	
Formato da reação	Reação homogénea		X		X	X	
	Reação com suporte sólido	X		X			X
Mecanismo de detecção	Deteção colorimétrica	X					
	Deteção de fluorescência		X		X		X
	FRET		X				
	Deteção de luminescência			X			
	Espectrometria de massa					X	

Relativamente o mecanismo de detecção utilizado na pirosequenciação (luminescência), trata-se de um mecanismo que recorre a enzimas que, na presença de pirofosfato libertado aquando da incorporação de nucleótidos na cadeia de DNA, ativam as propriedades bioluminescentes da luciferina, sendo essa luminescência monitorizada em tempo real. Por último, a espectrometria de massa é um método de detecção que tem por base o peso molecular do produto formado não recorrendo a qualquer tipo de marcação.

Segue-se a descrição individual dos métodos referidos e respetivas vantagens e desvantagens em termos de aplicação em análises forenses.

## 2.1. REVERSE DOT BLOTS

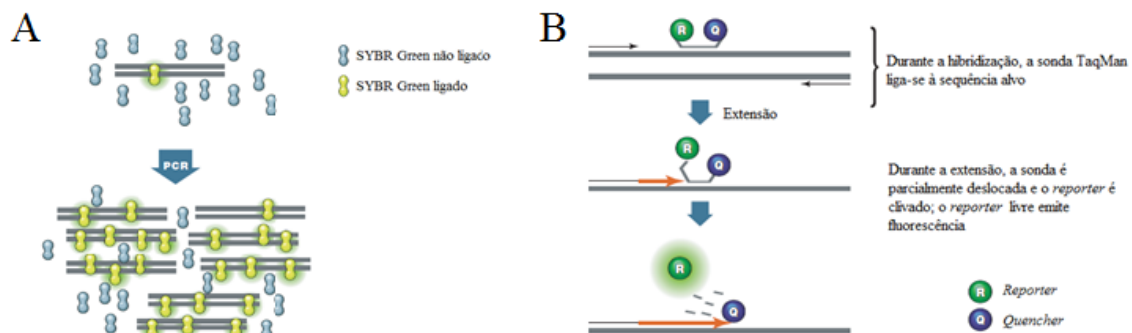
SOBRINO et al. (2005) refere este método como sendo o primeiro aplicado na detecção de polimorfismos analisados por PCR na área forense. Neste método de detecção as sondas alelo-específicas encontram-se ligadas a uma membrana de nylon, sendo que os diversos alelos do mesmo locus encontram-se representados uns imediatamente ao lado dos outros com diferentes sondas. Os produtos de PCR (sequência alvo de DNA) marcados com biotina hibridizam com as sondas correspondentes na membrana sob condições de temperatura controladas. O complexo streptavidina-peroxidase liga-se à biotina dos produtos de PCR que hibridizaram com as sondas e a posterior inclusão de um substrato cromogénico – a Tetrametilbenzidina ou TMB – que forma um produto azul quando exposto à ação da peroxidase do complexo referido permitem aferir a presença ou ausência de um determinado alelo na sequência alvo de DNA.

Exemplos de kits forenses que fazem uso deste método de detecção são o DQ Alpha AmpliType Kit lançado em 1990 pela Cetus Corporation (posteriormente adquirida pela Perkin-Elmer/Roche) e o AmpliType PM+DQA1 também conhecido como Poly-Marker (PM) que amplifica uma porção do gene HLA DQ juntamente com outros cinco segmentos de DNA (BUTLER, 2009). O último kit permite apenas um poder de discriminação de 1 em 10.000, o que fez com que este método viesse a ser posteriormente substituído pela genotipagem de STRs no campo da identificação. (BUTLER, 2009, SOBRINO et al., 2005). Relativamente às vantagens, (BUTLER, 2009) refere que se trata de um método rápido e

simples que não implica a utilização de equipamentos específicos (após a PCR) e útil em análises de amostras degradadas por fazer uso da PCR.

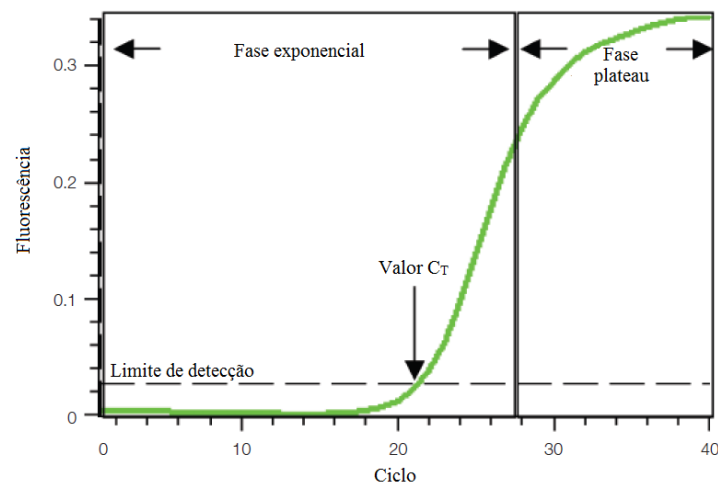
## 2.2. PCR QUANTITATIVA EM TEMPO REAL

A metodologia de PCR quantitativa em tempo real permite quantificar os produtos de amplificação da reação de polimerização (PCR) à medida que vão sendo gerados a cada ciclo, ou seja, “em tempo real”. A detecção dos produtos de PCR é realizada através da quantificação da emissão de fluorescência de fluoróforos que são incluídos na reação, sendo a quantidade de fluorescência emitida diretamente proporcional à quantidade de DNA. Existem vários sistemas de incorporação de fluoróforos, sendo que essas moléculas podem ser intercaladas nas cadeias recém formadas de forma inespecífica ou integrarem sondas ou primers específicos (como é o caso do corante SYBR GREEN I e das sondas TaqMan, respectivamente - Figura 25) (BIO-RAD LABORATORIES, 2006).



**Figura 25 – Esquemas ilustrativos de dois sistemas de incorporação de fluoróforos em reações PCR em tempo real. A – SYBR Green: os marcadores fluorescentes são incorporados de forma inespecífica nas cadeias de DNA recém formadas. B – sondas TaqMan: a sonda hibridiza com a sequência alvo e, entretanto o primer hibridiza a montante desse locus; a reação de polimerização tem início de seguida e, quando a polimerase atinge a extremidade 5’ da sonda a primeira leva a cabo a clivagem da segunda, libertando o fluoróforo, que, livre da ação do quencher, emite fluorescência; as sondas que não correspondem na totalidade à sequência alvo são deslocadas sem que ocorra a clivagem o que faz com que não haja emissão de fluorescência. Adaptado de BIO-RAD LABORATORIES (2006).**

A fluorescência emitida pelos fluoróforos é detetada por módulos de detecção de fluorescência incluídos em termocicladores especialmente desenhados para este tipo de reação que monitorizam a reação em tempo real, obtendo-se um gráfico de amplificação exemplificado na Figura 26. Nesse gráfico identificam-se duas fases: a fase exponencial na qual a quantidade de DNA praticamente duplica a cada ciclo e a fase plateau na qual o aumento da quantidade de DNA deixa de ser tão marcado devido ao facto de não haver tanta disponibilidade de componentes da reação.



**Figura 26 - Gráfico de amplificação de PCR em tempo real. Adaptado de BIO-RAD LABORATORIES (2006).**

Em termos de aplicação na genotipagem de SNPs, a PCR em tempo real é tida como um sistema de análise rápido, robusto, preciso e sensível que permite até detetar proporções muito pequenas de DNA masculino em material feminino (LAREU et al., 2001). Para além disso, permite a genotipagem de mais do que um SNP em cada reação, bastando para isso utilizar diferentes marcadores fluorescentes nas sondas (SOBRINO et al., 2005). O número de SNPs que podem ser genotipados em cada reação é, no entanto limitado ao número de marcadores fluorescentes que é possível utilizar (BÖLL et al., 2004, LAREU et al., 2001), o que constitui uma incontornável limitação.

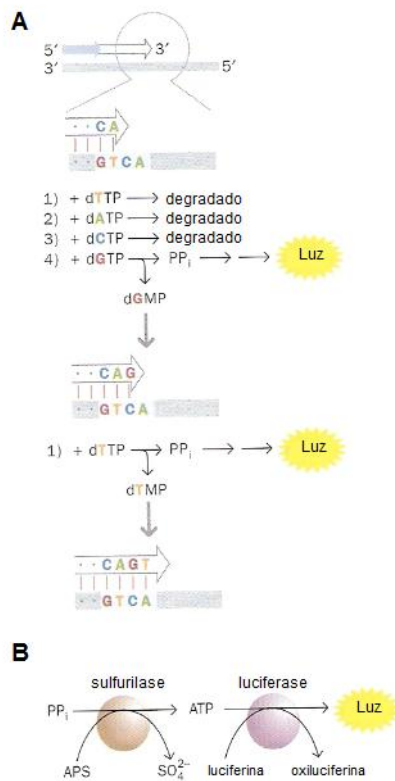
Relativamente às sondas TaqMan em particular, LIVAK (1999) refere que o seu uso permite obter um elevado rendimento que constitui uma mais-valia na genotipagem em larga escala de polimorfismos conhecidos. Por outro lado, e ainda segundo o autor, trata-se de um método que possui muitas vantagens em termos de tempo, exigência técnica e

custo por não haver necessidade de processamento pós-PCR. Por seu turno, NIEDERSTÄTTER et al. (2006) realça a possibilidade de quantificar proporções de alelos (fator importante na análise de DNA mitocondrial, nomeadamente no que toca à deteção de mistura ou heteroplasmia) entre outros aspetos que, segundo os autores, fazem deste método uma importante ferramenta que complementa métodos de elevada capacidade de teste simultâneo de vários marcadores (multiplex) como, por exemplo, o *SnaPshot*.

### 2.3. PIROSSEQUENCIAÇÃO

A pirosequenciação corresponde a uma tecnologia que tem por base o princípio da sequenciação por síntese, sendo esse mesmo processo monitorizado em tempo real por bioluminescência (Figura 27). O sistema de deteção tem por base a libertação de pirofosfato que ocorre aquando da inclusão de um (ou mais) nucleótido(s) na cadeia de DNA. Quatro enzimas intervêm no processo: a Klenow DNA polimerase (que incorpora os dNTPs na cadeia de DNA), a ATP sulfúrilase, a Luciferase e a Apirase. A ATP fosforilase converte o pirofosfato em ATP na presença de adenosina 5'fosfossulfato; de seguida, o ATP intervém na reação na qual a luciferase converte a luciferina em oxiluciferina, produto de reação com propriedades bioluminescentes, sendo que a quantidade de luz emitida é proporcional à quantidade de ATP; o sinal é detetado e registado por uma câmara e os dNTPs e ATP em excesso são degradados pela apirase. O processo é cíclico sendo que os dNTPs são introduzidos na reação de forma individual e sequencial (AHMADIAN et al., 2006, STRACHAN et al., 2010).

O método de sequenciação por pirosequenciação é altamente sensível, permitindo a obtenção de resultados a partir de apenas 8pg de DNA, o que o torna mais sensível do que o *SNaPshot*<sup>TM</sup>, MALDI TOF MS ou as sondas TaqMan, o que faz deste método uma valiosa ferramenta de genotipagem de SNPs no campo forense (HARRISON et al., 2006).

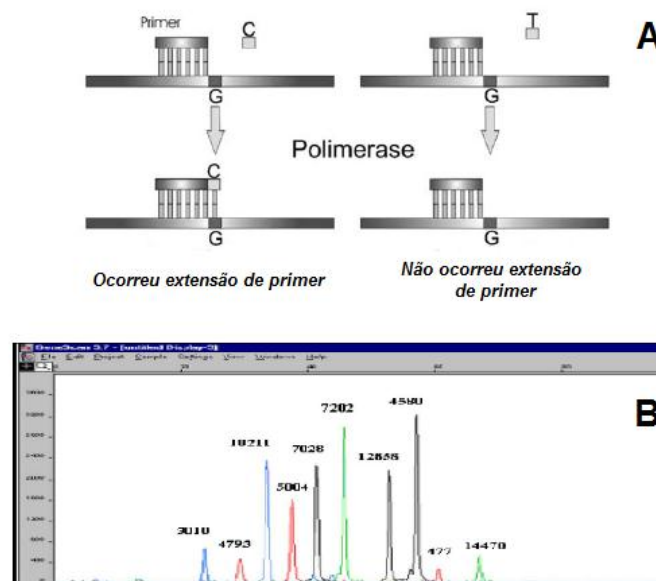


**Figura 27** – Esquema representativo do princípio do método de pirosequenciação. A – Síntese de DNA e simultânea identificação dos nucleótidos inseridos; aquando da síntese de DNA, a DNA polimerase adiciona dNTPs à nova cadeia de forma a complementar à cadeia molde; os dNTPs são adicionados à reação de forma individual e sequencial e quando é incorporado um dNTP transforma-se em dNMP e liberta pirofosfato que é usado nas reações posteriores para produzir luz; se o dNTP ministrado não for o complementar à cadeia molde, este é degradado pela apirase; na imagem, o primeiro dNMP incorporado é o G, como indicado pela produção de luz apenas nessa reação, seguindo-se o T. B – a inserção do dNMP correto é monitorizada pela produção de luz numa reação a dois tempos: o pirofosfato libertado é usado pela ATP sulfurilase para gerar ATP que, por sua vez leva a luciferase a transformar a luciferina em oxiluciferina que emite luz, luz essa que é detetada e registada pelo aparelho. Adaptado de STRACHAN et al. (2010).

## 2.4. MINISEQUENCIAÇÃO

O método de minissequenciação, primeiramente descrito por TULLY et al. (1996), implica a amplificação da sequência alvo a partir do DNA genómico e posterior purificação do DNA amplificado. O primer hibridiza com o produto de amplificação imediatamente a montante do local do SNP. Posteriormente, na presença dos quatro didesoxinucleótidos (ddNTPs) marcados com diferentes fluoróforos, ocorre a extensão de uma única base (PATI et al., 2004), como se encontra esquematizado na Figura 28 A. O produto desta reação é submetido a eletroforese capilar o que permite identificar qual o ddNTP incorporado e, consequentemente, genotipar o polimorfismo presente (Figura 28 B). No sentido de separar os produtos de reação espacialmente durante a eletroforese, podem ser adicionadas caudas de mobilidade modificadas à extremidade 5' de cada primer (BUTLER, 2009, TULLY et al., 1996).



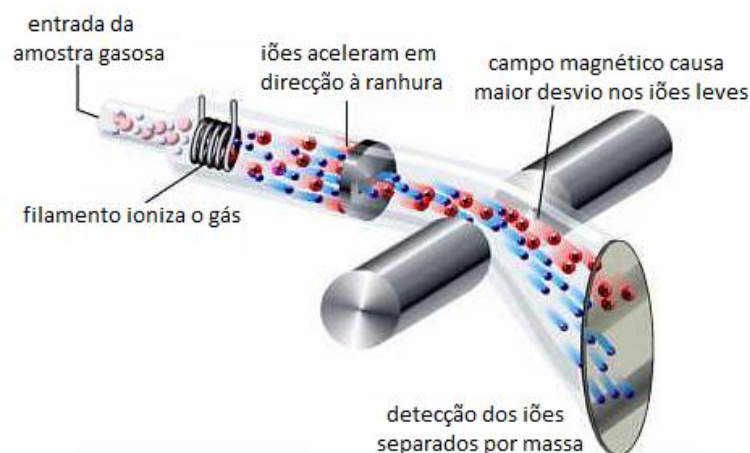


**Figura 28** – Esquema representativo da minissequenciação. A – Um primer liga-se à sequência alvo de DNA imediatamente a montante do SNP e é realizada a extensão do mesmo com um único nucleótido complementar à base polimórfica; adaptado de SOBRINO et al. (2005). B – Eletroferograma no qual estão representadas as amplificações de 10 SNPs de mtDNA; adaptado de ARGAC et al. (2009).

Trata-se de um método com elevada precisão que permite a genotipagem de até 30 SNPs em simultâneo (BULBUL et al., 2009) e que provou ser eficiente na genotipagem de amostras forenses, inclusivamente em amostras degradadas relativamente às quais não tinha sido possível a genotipagem de STRs (MOSQUERA-MIGUEL et al., 2009). Por outro lado, os equipamentos requeridos na execução deste método são os mesmos da genotipagem de STRs – termociclador para a reação de PCR e sequenciador automático para a posterior deteção e identificação, pelo que os laboratórios que já possuam estes equipamentos não necessitam de investir noutros para a realização da genotipagem de SNPs (BUDOWLE et al., 2009). Encontra-se disponível no mercado um sistema multiplex – SNaPshot® – comercializado pela Applied Biosystems que tem por base o método de minissequenciação. Várias são as publicações na área forense que fazem uso deste sistema, em particular na genotipagem de SNPs do mtDNA (ARGAC et al., 2009, MOSQUERA-MIGUEL et al., 2009, QUINTÁNS et al., 2004) e NRY (BENDER et al., 2003, CAEIRO et al., 2004).

## 2.5. ESPECTROMETRIA DE MASSA

A espectrometria de massa é uma tecnologia que permite a determinação precisa da massa de moléculas numa amostra. O espectrómetro de massa (Figura 29) é constituído por três componentes: o ionizador, o analisador de massa e o detetor de iões. O ionizador converte a amostra em iões em fase gasosa e provoca a sua aceleração em direção ao analisador de massa que, por sua vez separa os iões em função da relação massa/carga em direção ao detetor de iões. No detetor de iões ocorre o registo do impacto individual dos iões que permite obter o espectro de massa da amostra (STRACHAN et al., 2010).

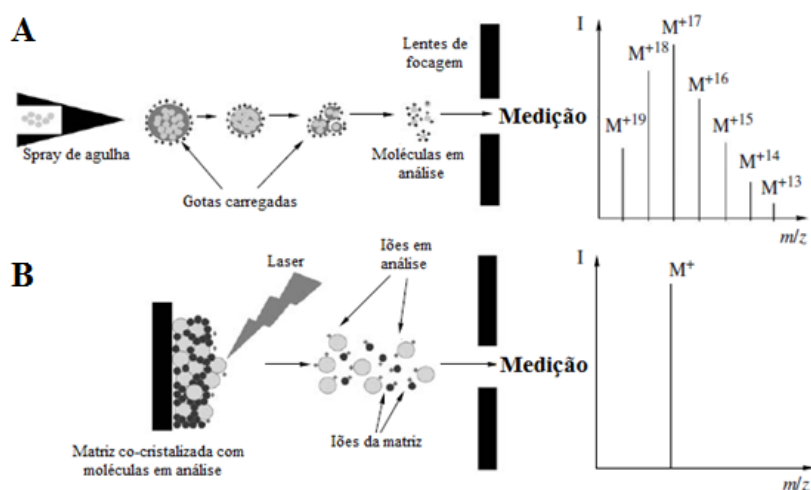


**Figura 29 - Esquema ilustrativo do funcionamento de um espectrómetro de massa.**  
Adaptado de: <http://www.dnassequencing.com/tag/maldi-tof-mass-spectrometry/>,  
consultado a 28-06-2011.

O processo de ionização provoca a fragmentação aleatória de grandes moléculas, como o DNA, pelo que se tornou necessário o desenvolvimento de métodos que permitissem a análise dessas moléculas sem que ocorresse essa mesma fragmentação – a ionização suave. Destacam-se, nesta matéria, dois métodos com ampla utilização em estudos moleculares: *Matrix-Assisted Laser Desorption/Ionization* (MALDI) e *Electrospray Ionization* (ESI). No primeiro método (Figura 30 B) a amostra é misturada com uma matriz com a capacidade de absorção de luz num solvente orgânico, solvente este que, ao evaporar, leva à formação de cristais de amostra e matriz que são aquecidos por laser. A dessorção térmica provocada pela energia do laser leva à expansão da matriz e amostra à

fase gasosa, sendo a amostra posteriormente ionizada e acelerada em direção ao detetor de massa (STRACHAN et al., 2010). A eficiência deste método depende, segundo IL'INA et al. (2009), de uma série de fatores, nomeadamente o método de preparação da amostra, o grau de purificação da mesma, o rácio matriz/amostra, a seleção adequada da substância da matriz, o poder de radiação do laser e a pressão de ionização. O método de deteção mais usado em conjunto com o método MALDI é o *Time of Light* (TOF), metodologia que tem por base o facto de a massa dos iões acelerados em análise ser inversamente proporcional à velocidade. A metodologia MALDI-TOF permite a análise de fragmentos de ácidos nucleicos de cadeia simples e dupla e outras biomoléculas até 100 kDa (IL'INA et al., 2009).

No método ESI (Figura 30 A), a amostra é dissolvida e a formação de iões deve-se à nebulização contínua dessa mesma solução quando esta é impelida através de um capilar na qual é exposta a elevada voltagem elétrica (2,5 a 4 kV). À saída do capilar, a solução dá origem a pequenas gotas com superfície fortemente carregada. Estas gotas perdem, posteriormente volume à medida que o solvente evapora e se movem em direção ao contraelétrodo. Estando sujeitas a elevada tensão, estas pequenas gotas sofrem fragmentações até que cada uma delas contém apenas uma partícula carregada da amostra que se pretende analisar. Os iões produzidos são, posteriormente detetados pelo aparelho (IL'INA et al., 2009)



**Figura 30 – Representação esquemática da produção de partículas com carga e posterior representação gráfica dos resultados obtidos através de dois métodos de espectrometria de massa: A - ESI; B - MALDI. Adaptado de IL'INA et al. (2009).**

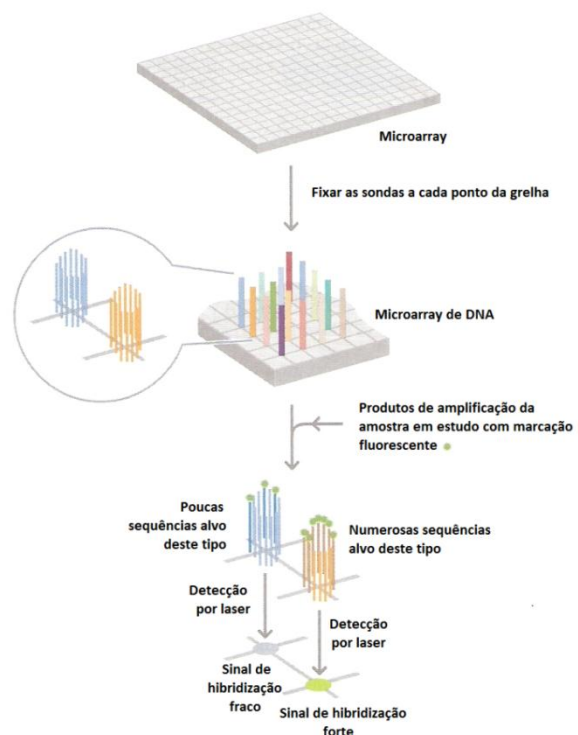
O método MALDI apresenta como vantagens a elevada rapidez, precisão e sensibilidade. No entanto implica um elevado investimento em equipamento e a execução de procedimentos complexos (em vários passos) que incluem purificações (PETKOVSKI et al., 2005).

A espectrometria de massa pode também ser utilizada em conjunto com a minissequenciação na análise de SNPs. LESSIG et al. (2004) fazem referência ao facto de a genotipagem de SNPs ser mais rápida quando é utilizado o método de detecção MALDI-TOF do que a eletroforese capilar, referindo ainda que, no entanto, o segundo método apresenta a vantagem de poder ser utilizado para a genotipagem de STRs.

## 2.6. MICROARRAYS DE DNA

Os microarrays de DNA (Figura 31), também conhecidos como chips de DNA ou biochips, correspondem a superfícies sólidas de vidro, plástico ou silicone nas quais se encontram fixados oligonucleótidos (sondas) num formato de grelha de alta densidade (STRACHAN et al., 2010). O número de marcadores presentes em cada microarray pode variar entre os milhares e milhões (a Affimetrix (<http://www.affymetrix.com/>), por exemplo, disponibiliza microarrays com 1.500 a 2,6 milhões de marcadores). Cada ponto da grelha possui milhares de cópias idênticas de um único tipo de sonda. Os produtos de amplificação da amostra marcados com fluoróforos são posteriormente adicionados, ocorrendo, de seguida, a hibridização entre estes e as sondas complementares. O sinal de hibridização (fluorescência) é posteriormente detetado, à semelhança de outros métodos que recorrem a marcação fluorescente, por laser. A intensidade desse mesmo sinal varia em proporção direta à quantidade de produtos de amplificação que hibridizaram com as sondas presentes no microarray.

**Figura 31 – Representação esquemática do princípio de hibridização de microarray. O microarray corresponde a uma superfície sólida à qual são fixadas sondas em coordenadas específicas com as quais os produtos de amplificação da amostra posteriormente hibridizam caso haja complementaridade entre estes e as sondas. Após a lavagem (que retira da reação os produtos não hibridizados) é detetada, por laser, a fluorescência emitida pelos produtos de amplificação hibridizados com as sondas que permanecem fixos ao microarray o que permite identificar a(s) sequência(s) presente(s) na amostra. Adaptado de STRACHAN et al. (2010).**



O sinal emitido em cada ponto da grelha é posteriormente analisado recorrendo a software adequado que permite a conversão desse mesmo sinal numa paleta de cores de acordo com a intensidade do mesmo (STRACHAN et al., 2010).

Existem também métodos que combinam o princípio da hibridização subjacente à tecnologia do microarray e a extensão de primer (BOGUS et al., 2006, PASTINEN et al., 2000, SOBRINO et al., 2004). Esta mesma combinação de princípios é aplicada em dois inovadores sistemas – GeneChip da Affimetrix™ e BeadArray da Illumina™ (STRACHAN et al., 2010). Estas tecnologias permitem a genotipagem de milhares a milhões de SNPs em simultâneo, o que, segundo SOBRINO et al. (2005) corresponde a um número excessivo relativamente às necessidades atuais dos laboratórios forenses.

Em suma, as principais vantagens dos microarrays são o elevado paralelismo e automatização e as suas reduzidas dimensões o que permite a análise rápida de uma enorme quantidade de dados relativos a vários loci em simultâneo (TENG et al., 2009). Os microarrays de DNA podem ser utilizados em diversas aplicações desde a genotipagem forense de SNPs (KRJUTSKOV et al., 2009, LI et al., 2006, SOBRINO et al., 2004) ao

mapeamento de características complexas através de estudos que envolvem todo o genoma (GWAS – *genome wide association studies*).

## 2.7. MÉTODOS DE SEQUENCIAÇÃO DE PRÓXIMA GERAÇÃO E PERSPETIVAS FUTURAS

Recentemente têm surgido no mercado novas tecnologias de sequenciação designadas tecnologias de “sequenciação da próxima geração” (NGS – *next-generation sequencing*) que têm por base a fragmentação do genoma em pequenos segmentos que são posteriormente ligados a adaptadores para leitura aleatória durante a síntese de DNA (sequenciação por síntese) (ZHANG et al., 2011). Cinco plataformas atualmente disponíveis e um conjunto de outras vinte e uma em desenvolvimento são referidas por ZHANG et al. (2011), que afirma, em relação ao primeiro conjunto, que as plataformas Roche/454 FLX, the Illumina/Solexa Genome Analyzer e (ABI) SOLiD Analyzer dominam atualmente o mercado, enquanto a Polonator G.007 e Helicos HeliScope são as mais recentes, sendo que o seu uso não é ainda disseminado. O autor enfatiza a importância da elevada capacidade de obtenção de informação com base nas tecnologias NGS no âmbito dos estudos de associação que permitem a identificação das bases genéticas de doenças ou outros traços complexos.

Para além da descoberta de variantes genéticas por re-sequenciação de regiões alvo de interesse ou até de genomas inteiros, METZKER (2010) aponta ainda outra aplicação destas tecnologias: a construção de perfis globais de marcadores epigenéticos e estrutura da cromatina, citando WANG et al. (2009) e PARK (2009). Em concordância com o autor anteriormente referido, MARDIS (2008) acrescenta ainda que estas plataformas permitem a determinação dos padrões de metilação em todo o genoma e da forma como esses padrões se alteram durante o desenvolvimento de um organismo, quer no contexto de uma doença, quer sob a influência de outros fatores.

As tecnologias NGS estão também a ser usadas no âmbito da caracterização de relações evolutivas entre genomas ancestrais, destacando-se a aplicação na análise de genomas de Neandertais (BRIGGS et al., 2009, GREEN et al., 2006).

No que diz respeito à potencial aplicação das tecnologias NGS na rotina forense, IRWIN et al. (2011) expressam otimismo, afirmando que é provável que estas tecnologias venham a permitir a construção de perfis completos do genoma mitocondrial até mesmo a partir de amostras degradadas de forma simples e com baixo custo. Por seu turno, BANDELT et al. (2011) não se mostra tão otimista, afirmando, com base nos trabalhos de HE et al. (2010) e ZARAGOZA et al. (2010) que os dados obtidos através da aplicação das tecnologias NGS não respeitam os critérios de qualidade exigidos na análise de amostras forenses. BANDELT et al. (2011) enfatiza a inexistência de avaliações reportadas relativamente à qualidade dos dados obtidos, defendendo que, para solucionar o problema, é necessária a realização prévia de exercícios de uniformização antes da implementação dessas tecnologias na prática forense. Os autores referem que a elevada quantidade e qualidade de DNA requerida pelos métodos NGS é outra desvantagem, o que contraria o que foi referido por IRWIN et al. (2011).





## CAPÍTULO 3

### 3. APLICAÇÕES PRÁTICAS DOS MARCADORES GENÉTICOS EM CONTEXTO FORENSE

Quando um perfil de DNA de uma amostra colhida numa cena de crime não encontra correspondência quer em amostras referência de suspeitos quer numa base de dados, qualquer informação que possa ser deduzida a partir do DNA da amostra sobre o seu dador é útil (JOBLING et al., 2004). Atualmente existem duas áreas de previsão de fenótipos – indireta (via ancestralidade biogeográfica) e direta (estudo das determinantes genéticas de características fenotípicas) – que apresentam um grande potencial de utilização na área da genética forense e, por isso, têm sido alvo de investigação (JOBLING et al., 2004). Estas áreas permitem obter informações acerca da aparência do dador da amostra que se pretende identificar, funcionando como uma espécie de testemunha ocular.

#### 3.1. PREVISÃO DE FENÓTIPOS – MÉTODO DIRETO

O método direto de previsão de características fenotípicas consiste na identificação dos elementos genéticos subjacentes a essas mesmas características no sentido de tornar

possível a previsão da probabilidade de que num indivíduo se observe um determinado fenótipo com base na análise do seu DNA. Este método é mais vantajoso do que o método indireto por fornecer informação com melhor qualidade, no entanto não é muitas vezes aplicável, sendo também mais dispendioso e complexo em termos de aplicação (FRUDAKIS, 2008). Estas desvantagens estão diretamente relacionadas com a natureza complexa dos traços em abordagem. De facto, como foi já referido e de acordo com PULKER et al. (2007), os traços complexos são determinados por um conjunto de genes de número variável com características particulares. Estas características dificultam a identificação desses traços complexos. No entanto, nos últimos anos têm vindo a ser criadas bases de dados com milhares de marcadores de diversas populações (Human Genome Diversity Cell Line Panel – HGDP-CEPH (CAVALLI-SFORZA, 2005), *the allele frequency database* – ALFRED (<http://alfred.med.yale.edu/>), *National Centre for Biotechnology Information* (NCBI) dbSNP e HapMap (PHILLIPS, 2007), por exemplo), o que, em paralelo com o número crescente de estudos de associação, tem permitido a crescente acumulação de conhecimentos relativos às determinantes genéticas dos traços complexos (MCCARTHY et al., 2008, PULKER et al., 2007).

No sentido de estabelecer a relação entre o fenótipo e as bases genéticas que lhe estão subjacentes é necessária a caracterização dos traços fenotípicos (fenotipagem) que, em alguns casos (como a altura) é bastante simples, mas noutros pode tornar-se bastante complexa como é o caso dos traços relacionados com a pigmentação – cor dos olhos, cabelo e pele. Neste sentido, em relação aos traços relacionados com a pigmentação, diferentes autores recorrem a diferentes abordagens relativamente aos procedimentos de fenotipagem. O grupo de Frudakis, por exemplo, privilegia os métodos espectrofotométricos e análise de fotografia digital em ambiente controlado por considerar que a avaliação subjetiva constitui uma fonte de erro. No trabalho do autor intitulado “Molecular Photofitting” encontram-se descritos vários métodos de fenotipagem que classificam de forma objetiva características fenotípicas como o índice intercantal (para descrever a forma dos olhos em vez de “olhos com dobras epicânticas”) e o índice de melanina (Índice M) em vez de “tom de pele” (FRUDAKIS, 2008). Por seu turno, o grupo de Kayser recorre normalmente a métodos de avaliação por observação (auto e heterodescrição), uma vez que, segundo BRANICKI et al. (2011) com base em estudo anteriores, que não existem evidências de que os métodos de avaliação quantitativa sejam melhores do que os métodos baseados em auto e heterodescrição.

Outro aspeto a ter em consideração relativamente a características como o tom de pele, o índice de massa corporal e a estatura é o facto de estas se encontrarem sujeitas à ação de fatores ambientais (FRUDAKIS, 2008, KAYSER et al., 2009). Nesses casos há que ter uma especial atenção na determinação do método de classificação fenotípica pois haverá a necessidade de discernir entre a influência dos fatores genéticos e dos fatores ambientais (FRUDAKIS, 2008). Por outro lado, algumas características físicas podem ser alteradas de forma voluntária: o cabelo pode ser pintado, a cor dos olhos pode ser dissimulada através do uso de lentes de contacto ou até, como tem sido recentemente noticiado (BBC NEWS, 2011, LIMER, 2011), alterado de forma permanente através de cirurgia laser, existindo inclusivamente uma patente registada nos Estados Unidos da América desde 2001 intitulada “Métodos de alteração da pigmentação da íris” (HOMER, 2001).

### *3.1.1. GÉNERO OU SEXO*

O género ou sexo é uma característica cuja determinação através da análise de DNA é possível há mais de quinze anos (MANNUCCI et al., 1994) através do estudo do gene da Amelogenina que apresenta cópias com diferentes tamanhos nos cromossomas X e Y, diferença essa que origina fragmentos de tamanho diferente no caso de a amostra ser de um indivíduo do sexo masculino. Trata-se de um marcador que se encontra incluído na maioria dos testes de identificação forense, não estando, no entanto, livre de erro. Estudos realizados em diferentes populações apontam para uma taxa de erro que pode variar entre os 0,018% (2 de 29.432 indivíduos) na população Austríaca (STEINLECHNER et al., 2002) e os 8% (2 de 24 indivíduos) no Sri Lanka (SANTOS et al., 1998). Na população Indiana foram reportadas taxas de erro de 1,85% (5 em 270 indivíduos) (THANGARAJ et al., 2002) e 0,23% (10 em 4.257 indivíduos) (KASHYAP et al., 2006), na população de Israel foi identificado um caso em 96 amostras, o que representa uma taxa de erro de 1,04% (MICHAEL et al., 2004) e na população do Nepal foi apurada uma taxa de erro de 6,49% (5 em 77 casos analisados) (CADENAS et al., 2007). MACIEJEWSKA et al. (2009) e SHADRACH et al. (2004) apontam para que os erros na determinação do sexo do indivíduo se devam à ocorrência de uma deleção do cromossoma Y ou a uma mutação pontual na região do cromossoma X à qual se liga o primer reverso do gene da Amelogenina em vários kits

comerciais de identificação. O facto de as taxas de erro mais elevadas terem sido reportadas em países do sul da Ásia, segundo CADENAS et al. (2007) não é obra do acaso. Os autores referem que as deleções da amelogenina observadas no cromossoma Y são mais prevalentes numa região específica – o sub-continente Indiano, estando associadas a um haplogrupo específico (J2b2-M241) nos cinco indivíduos analisados no referido trabalho, o que se coaduna com a correlação existente entre a estrutura filogenética do cromossoma Y e a geografia. CADENAS et al. (2007) afirmam, porém, que serão necessários mais estudos relativos a vários tipos de marcadores e outras amostras de indivíduos com a mutação para perceber se as deleções observadas noutros estudos surgiram através da ocorrências de mutações independentes ou se estão relacionadas por descendência (se se observa o mesmo haplogrupo).

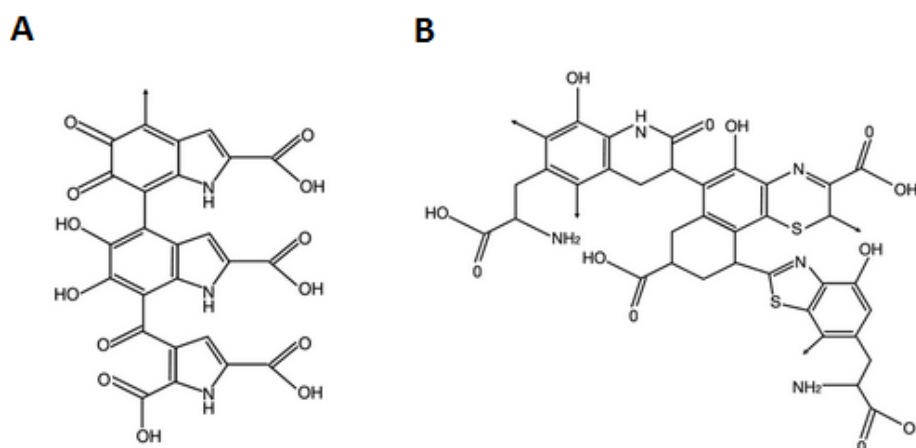
Assim, é desejável não só que os resultados sejam interpretados de forma cuidadosa, mas também que sejam adicionados aos testes genéticos outros marcadores de forma a evitar esses mesmos erros, em particular marcadores STR do cromossoma Y (MICHAEL et al., 2004) ou outros marcadores cuja localização seja distante da amelogenina (KAYSER et al., 2009).

Por outro lado, VON WURMB-SCHWARK et al. (2007) refere casos nos quais se observam fenómenos de quimerismo (transplantes de medula óssea em que o dador e recetor têm diferentes géneros (DAUBER et al., 2004)) ou micro-quimerismo (mulher grávida cujo feto é do sexo masculino (BIANCHI, 2004)) representam também um complexo desafio em termos de interpretação. Os mesmos autores discutem também a importância da discrepância entre o género biológico e o género legal presente nos documentos identificativos observada em casos de intersexualidade e transexualidade. Nesses casos, o fenótipo observado e/ou os registos legais podem indicar um género contrário ao biologicamente testado, facto que, segundo VON WURMB-SCHWARK et al. (2007) não deve ser menosprezado.

### *3.1.2. TRAÇOS RELACIONADOS COM A PIGMENTAÇÃO*

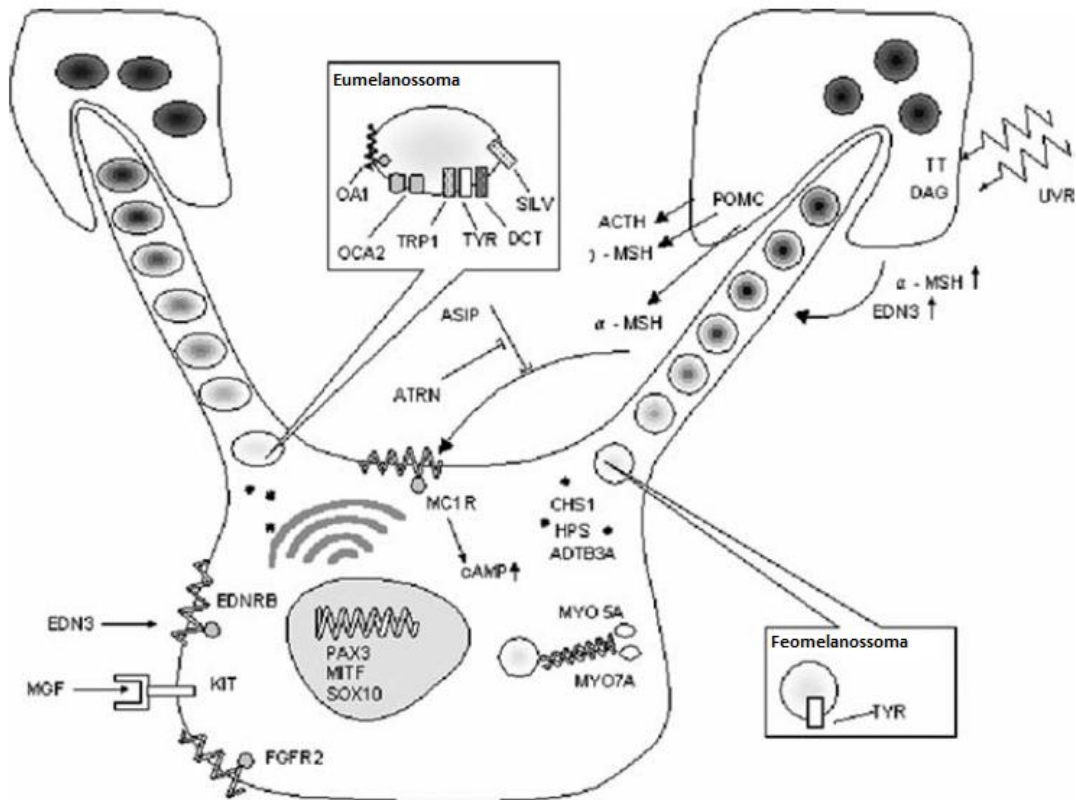
Sendo uma característica com forte componente hereditária e facilmente classificável, os traços relacionados com a pigmentação tornaram-se um alvo preferencial

de pesquisa genética ao longo das últimas décadas (FRUDAKIS, 2008). Traços como o tom de pele, a cor dos olhos e a cor do cabelo partilham o facto de serem influenciados principalmente pela quantidade e tipo de melanina presente nessas mesmas estruturas (sendo também influenciadas, embora em muito menor escala, pelas concentrações de outros pigmentos como a hemoglobina das células sanguíneas e carotenóides (REES, 2003)). A melanina corresponde a uma molécula que tem a capacidade de absorver a luz e que possui uma grande resistência à degradação química (STURM et al., 2004). Existem dois tipos de pigmento melânico: a eumelanina (castanha/preta) e a feomelanina (amarela/vermelha) sendo que as diferenças entre estes dois tipos assentam na estrutura molecular (Figura 32) e, segundo VALENZUELA et al. (2010), no teor de enxofre.



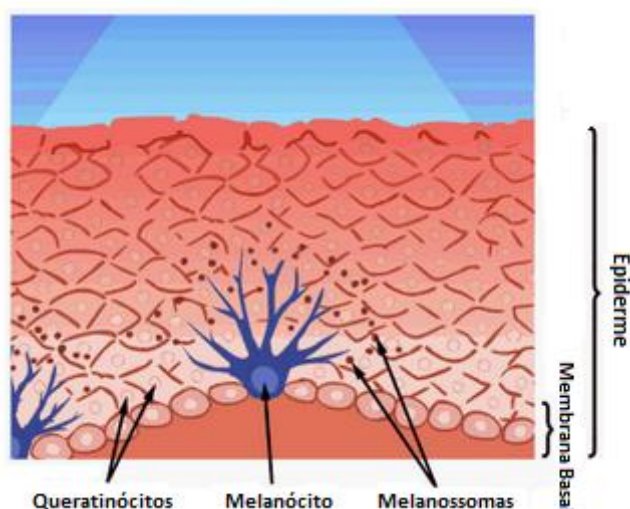
**Figura 32 – Estrutura molecular dos pigmentos melânicos: A – Eumelanina; B – Feomelanina.** Adaptado de : <http://palaeo.gly.bris.ac.uk/melanosomes/melanin.html>, consultado a 12-05-2011.

O processo de síntese de melanina (melanogénese) tem lugar no melanossoma, organelo citoplasmático dos melanócitos (Figura 33) que são as células produtoras do pigmento (JUNQUEIRA et al., 2004).



**Figura 33 – Representação esquemática de um melanócito e vias de síntese de melanina.** Após a transcrição no núcleo e tradução no citoplasma, o pigmento eumelânico ou feomelânico é depositado nos melanossomas (eumelanossomas ou feomelanossomas, respetivamente). A maturação destes organelos é representada pelo escurecimento sucessivo observado no sentido ascendente nas dendrites do melanócito. No esquema encontram-se representados vários genes envolvidos nas duas vias de pigmentação referidas. UVR – radiação ultravioleta. Fonte: FRUDAKIS (2008).

Na pele, os melanócitos localizam-se ao nível da membrana basal (Figura 34) e no cabelo, ao nível do bulbo capilar. Nestas estruturas os melanócitos surgem associados aos queratinócitos, células para as quais os melanossomas contendo melanina são transferidos ficando ali depositados (FRUDAKIS, 2008, JUNQUEIRA et al., 2004).



**Figura 34 - Representação esquemática da localização de um melanócito e outras estruturas intervenientes no processo de pigmentação da pele. Adaptado de:** <http://palaeo.gly.bris.ac.uk/melanosomes/melanin.html>, consultado a 12-05-2011.

A pigmentação da pele e cabelo é, então, influenciada pelo tamanho, número, distribuição e taxa de formação dos melanossomas bem como pelo tipo de melanina presente nos mesmos (FRUDAKIS, 2008). Relativamente à cor dos olhos, trata-se de uma característica determinada pelo tipo, distribuição e quantidade de melanina na íris que é uma estrutura que funciona como diafragma do olho controlando a quantidade de luz que atinge a retina (LIU et al., 2010).




Ao longo dos anos, vários genes cujos produtos intervêm no processo de melanogénese têm sido apontados como associados à diversidade fenotípica observada nos traços relacionados com a pigmentação (STURM, 2009). Segue-se uma descrição dos traços relacionados com a pigmentação (cor dos olhos e do cabelo e tom de pele) e do conhecimento atualmente disponível relativo à previsão dos mesmos.

### 3.1.2.1. COR DOS OLHOS

A cor dos olhos é, como foi anteriormente referido, uma característica dependente do tipo, distribuição e quantidade de melanina na íris (LIU et al., 2010). Assim, nos olhos castanhos observa-se uma maior quantidade de melanina presente nos melanossomas do que nos olhos verdes e azuis. As diferenças entre tons são consequência da concentração e

tipo de melanina presentes (Tabela 9) e não do número de melanócitos (relativamente ao qual não se observa variação). A eumelanina está presente em maioria nos olhos com tons que variam entre o castanho-escuro, cor de avelã e castanho-esverdeado. Já o tom verde é caracterizado pela presença de feomelanina em maior quantidade. Relativamente aos tons azuis, são determinados pela baixa concentração tanto de eumelanina como de feomelanina (FRUDAKIS, 2008, STURM et al., 2004).

**Tabela 9 – Relação entre a cor da íris e a concentração de eumelanina e feomelanina presentes nessa mesma estrutura.**

Cor da íris	Concentração	
	Eumelanina	Feomelanina
	+	-
	-	+
	-	-

Os sinais “+” e “-” referem-se a concentrações maiores e menores dos pigmentos relativamente às outras cores da íris representadas, respetivamente.

As imagens das cores da íris foram adaptadas de: [http://allwomenstalk.com/wp-content/uploads/eye\\_color.jpg](http://allwomenstalk.com/wp-content/uploads/eye_color.jpg), consultado a 13-05-2011.

Esta característica apresenta uma variação mais elevada em indivíduos com ancestralidade Europeia, estando relacionada com a latitude nesse mesmo continente (FROST, 2006, FRUDAKIS, 2008, STURM et al., 2004), como será abordado adiante. Por este motivo, o primeiro teste para previsão da cor dos olhos com aplicação forense, o RETINOME™ da DNAPrint Genomics, permitia, desde o seu lançamento em 2004, prever com 97% de precisão a cor castanha ou azul da íris apenas em indivíduos com proporção de ancestralidade Europeia superior a 80% (DNA PRINT GENOMICS, 2004). A mesma empresa lançou em 2006 uma nova versão deste teste (o RETINOME™ 2.0) que, segundo foi comunicado à imprensa (DNA PRINT GENOMICS, 2006) incluiu outros marcadores do gene *OCA2* (*oculocutaneous albinism II*, que codifica a proteína P) permitindo uma melhoria na precisão de previsão da cor dos olhos. Não consta, porém, no comunicado em que medida é melhorada a deteção e a restrição em termos de população alvo mantém-se.



Este teste foi também disponibilizado sob a forma de kit de eletroforese capilar permitindo a sua realização em laboratórios fora da empresa. É, no entanto de prever, que os marcadores presentes em ambos os testes sejam marcadores do OCA2, a julgar pelo trabalho publicado em 2008 por Frudakis – “Molecular Photofitting” (FRUDAKIS, 2008) e outros trabalhos do mesmo autor e colaboradores (DUFFY et al., 2007, FRUDAKIS et al., 2007, STURM et al., 2004).

Os referidos estudos focaram-se no OCA2 por ser um gene que, quando mutado, leva à ocorrência de uma forma de albinismo. No entanto, um estudo de DUFFY et al. (2007) revelou uma associação mais significativa apenas ao nível do primeiro intrão, o que sugere que o sinal de associação tem origem fora do gene OCA2. De facto, a variação no HERC2, gene com localização próxima do gene OCA2, apresenta uma associação mais forte à cor da íris do que a variação no OCA2, o que sugere que o HERC2 terá um efeito regulador do OCA2, influenciando a sua expressão, encontrando-se, por isso, na base da variação entre os fenótipos azul e castanho da cor da íris (EIBERG et al., 2008, KAYSER et al., 2008, KAYSER et al., 2009, STURM et al., 2008, SULEM et al., 2007).

Relativamente aos tons intermédios, estes constituem um grande desafio em termos de caracterização fenotípica, o que torna muito difícil a determinação das bases genéticas inerentes a esse mesmo grupo (FRUDAKIS, 2008, KAYSER et al., 2009). Esta dificuldade pode levar ao insucesso muitos estudos de associação se forem negligenciados fatores importantes como a presença de mais do que um tom e o padrão de distribuição de cor na íris. No entanto ter em conta estes fatores leva a um aumento exponencial do número de categorias de classificação e à necessidade de um elevado número de amostras para garantir a representatividade do maior número possível de fenótipos diferentes (FRUDAKIS, 2008). Assim, alguns autores defendem a aplicação de métodos como a avaliação do índice de melanina da íris com base na soma dos valores de luminosidade e cor (FRUDAKIS, 2008), quantificação da variação contínua da cor dos olhos em dois eixos correspondentes aos valores de tonalidade e saturação (LIU et al., 2010) por constituir uma abordagem que permite uma avaliação mais objetiva e abrangente do que a avaliação com base na observação direta (KAYSER et al., 2008), visto que a última permite apenas diferenciar azul, castanho e intermédio, classificação esta que não inclui grande parte da complexidade fenotípica, para além do facto de que alguns casos podem ser muito difíceis de classificar através de observação (Figura 35).



**Figura 35 – Figura ilustrativa da variação da cor da íris desde o azul ao castanho escuro passando por tons intermédios (cinza, verde, avelã e castanho claro). É possível identificar em algumas imagens pequenos pormenores como a presença de nevus, cristas, e outros como uma orla eumelânica que irradia a partir da pupila nas íris cor-de-avelã. Fonte: STURM et al. (2004).**

Por seu turno, MENGEL-FROM et al. (2010) reconhecem as vantagens da classificação por métodos de medição automatizados, afirmando que, para além do aumento da objetividade da classificação, é possível que se observem associações mais significativas quando estes métodos são usados. Os autores alegam, no entanto, que a aplicação da previsão da cor dos olhos na prática forense implica a classificação dessa característica com base na interpretação humana das cores e não em medições objetivas, pelo que a abordagem que recorre à autodescrição se torna interessante neste contexto.

Outros genes de pigmentação que têm vindo a ser estudados e apontados como muito influentes na variação da cor dos olhos são o SLC24A4, TYR (SULEM et al., 2007), SLC24A2 e IRF4 (LIU et al., 2009). Também os genes TYRP1 e ASIP foram analisados no contexto de previsão da cor da íris, tendo LIU et al. (2009) concluído que se trata de genes com pequeno efeito aditivo. Este mesmo estudo concluiu que, com base em apenas seis SNPs de seis genes (OCA2, HERC2, SLC24A4, SLC45A2, TYR e IRF4) é possível prever a variação fenotípica da cor da íris com 93, 91 e 72% de precisão para os fenótipos castanho, azul e intermédios, respetivamente, em indivíduos com ancestralidade Europeia. Um trabalho publicado recentemente (WALSH et al., 2011) reportou dados relativos aos mesmos seis marcadores, designando o conjunto por “sistema IrisPlex” refere que a

precisão da previsão deste conjunto de marcadores prever com mais de 90% de precisão os fenótipos azul e castanho relativos à cor dos olhos, sendo que a precisão de previsão para as cores castanha e azul com base apenas na análise do rs12913832 (no HERC2) ascende aos 89,9 e 87,7%, respetivamente. Este conjunto de marcadores permite, segundo os autores, a previsão do traço fenotípico independentemente da ancestralidade do dador da amostra (ao contrário do que se verifica relativamente ao RETINOME™) inclusivamente em amostras degradadas ou escassas por necessitar de apenas uma pequena quantidade de DNA (31 pg) para a realização do teste. Para além disso, o facto de este teste ser realizado com recurso à tecnologia SnaPshot e eletroforese capilar (tecnologias que se encontram implementadas na maioria dos laboratórios forenses) faz deste sistema de previsão uma ferramenta importante e facilmente implementável no âmbito da genética forense.

Ainda no âmbito da previsão da cor dos olhos, SPICHENOK et al. (2010) reportam também um conjunto de seis marcadores, três dos quais comuns ao sistema IrisPlex (SNPs com localização no HERC2, IRF4 e SLC45A2) e os restantes três localizados nos genes OCA2, MC1R e ASIP. A caracterização dos fenótipos foi levada a cabo através de autodescrição, tendo os indivíduos participantes no estudo preenchido um questionário que incluiu também informação sobre a ancestralidade biogeográfica. Os autores referem que a taxa de erro de previsão da cor dos olhos deste conjunto de marcadores é de 1% na população Europeia, não se tendo observado erros de previsão nas restantes populações testadas (Afro-Americana, Asiáticas (Sul e Oriental) e Mistura). Estes resultados devem-se, em parte, ao facto de a previsão em alguns casos ser realizada sob a forma de exclusão (ou seja, por exemplo, a previsão “não-castanho” inclui as cores azul e verde) o que diminui a taxa de erro, mas diminui também a precisão da previsão. Mais uma vez, o rs12913832 (no HERC2) é apontado como o principal elemento na previsão.

#### 3.1.2.2. COR DO CABELO

A cor do cabelo é uma característica com distribuição contínua embora seja possível a sua classificação em categorias nas quais são reconhecidas diferenças ao nível do tamanho e forma dos melanossomas bem como a deposição de melanina sua matriz

intracelular. No cabelo ruivo observa-se uma maior deposição de feomelanina relativamente à eumelanina em melanossomas pequenos e redondos embora seja também comum a coexistência de eumelanócitos parcialmente diferenciados contendo ambos os pigmentos na mesma célula. Há também casos de cabelo ruivo eumelânico, visto que, quando sofre oxidação (química – peroxidação – ou pela ação da radiação UV), a eumelanina adquire um tom amarelado/avermelhado semelhante ao observado na feomelanina. Assim, em certos casos, o fenótipo de cabelo ruivo é determinado pela presença de eumelanina que é naturalmente peroxidada, sendo, também esta, uma característica determinada geneticamente. Relativamente aos tons mais escuros (castanho e preto), observa-se a presença de melanossomas elípticos de grandes dimensões com grande densidade de eumelanina. O tom do cabelo torna-se tanto mais escuro quanto maior o tamanho dos melanossomas e a densidade do pigmento depositado. O cabelo loiro caracteriza-se pela presença de melanócitos em igual número mas com tamanho reduzido e menor pigmentação comparativamente a outros tons de cabelo. Por último, no cabelo branco observa-se uma diminuição acentuada do número de melanócitos, sendo que os melanócitos que persistem apresentam uma extensa redução em termos de conteúdo de melanina (FRUDAKIS, 2008).

Relativamente ao conhecimento das bases genéticas subjacentes à cor do cabelo, há um fenótipo que se destaca – cabelo ruivo. Há muito que é conhecida a associação entre este fenótipo e o gene MC1R (VALVERDE et al., 1995), associação essa que, segundo BRANICKI et al. (2011) e KAYSER et al. (2009) foi já confirmada em vários outros estudos, tendo sido esse conhecimento sido aplicado em análises forenses apesar das limitações em termos populacionais (trata-se de um fenótipo que está quase exclusivamente presente na população Europeia). Um dos estudos de âmbito forense reportou uma precisão de previsão do fenótipo de cabelo ruivo de 96% numa amostra de indivíduos Britânicos (GRIMES et al., 2001) recorrendo à genotipagem de 12 marcadores do MC1R. A este gene e fenótipo de cabelo ruivo estão também associadas a ocorrência de sardas e pele clara, sendo que os indivíduos com estas características apresentam um risco aumentado de desenvolver melanoma (PALMER et al., 2000, SULEM et al., 2008). SULEM et al. (2008) apontam o ASIP como outro gene associado ao fenótipo cabelo ruivo. Por outro lado foram identificados, em estudos GWA, SNPs nos genes MATP, OCA2, SLC24A4, HERC2, KITLG, TYR, TYRP1, TPCN2, IRF4, EXOC2, KIF26A, OBSCN como estando associados à determinação da cor do cabelo em Europeus (BRANICKI et al., 2011, HAN et al., 2008, KAYSER et al., 2009, SULEM et al., 2008, SULEM et al., 2007).

Um estudo recente (BRANICKI et al., 2011) que teve como amostra um grupo de 385 Polacos concluiu que é possível prever com elevada precisão a cor de cabelo com base no DNA. Recorrendo a um conjunto de 13 marcadores em 11 genes (MC1R, HERC2, IRF4, TYR, EXOC2, SLC45A2, TYRP1, OCA2, SLC24A4, KITLG e ASIP) os autores afirmam que a precisão de previsão atinge os 81% em relação ao fenótipo de cabelo loiro, 82% para o cabelo castanho, 87% para o cabelo preto e 93% relativamente ao fenótipo de cabelo ruivo (valores de precisão ajustada à prevalência). O trabalho incluiu também uma caracterização mais detalhada do fenótipo da cor do cabelo em 7 categorias (dividindo o fenótipo loiro em loiro e loiro-escuro e o fenótipo ruivo em castanho-avermelhado, loiro-avermelhado e ruivo) obtendo, ainda assim, bons resultados em termos de precisão de previsão (loiro - 78%; loiro-escuro - 73%; castanho - 82%; castanho-avermelhado - 82%; loiro-avermelhado - 92%; ruivo - 94%; preto - 88%). Os autores do estudo reconhecem, porém, que as limitações em termos de amostra (em termos de número e diversidade populacional) podem ter influenciado os resultados. Visto que a amostra foi “enriquecida” com indivíduos com fenótipo de cabelo ruivo, no sentido de demonstrar a precisão de previsão desse mesmo fenótipo, o que faz com que a amostra não seja representativa da população da Polónia, na qual este fenótipo é raro. Por outro lado o estudo não detetou associação entre um SNP (rs1426654) no gene SLC24A5 e a determinação da cor do cabelo, ao contrário do estudo de VALENZUELA et al. (2010) que reportou a associação entre este SNP e a variância fenotípica total da melanina no cabelo. Esta inconsistência dever-se-á, segundo BRANICKI et al. (2011), ao facto de este marcador refletir uma estrutura populacional subjacente à amostra do estudo de VALENZUELA et al. (2010), ou seja, este locus está associado à origem biogeográfica e não à característica fenotípica em questão. Em suma, o trabalho de BRANICKI et al. (2011) sugere que, com um pequeno número de marcadores, é possível prever a cor do cabelo de um indivíduo com base no seu DNA com elevada precisão, o que faz com que a previsão da cor do cabelo seja aplicável em análises forenses de previsão de características físicas.

Ainda em relação à cor do cabelo, há também que ter em atenção que, ao contrário da cor dos olhos, esta característica sofre alterações ao longo da vida, observando-se por exemplo, que indivíduos loiros durante a infância vêm a sua cor de cabelo escurecer na adolescência (REES, 2003).

### 3.1.2.3. TOM DE PELE

O tom de pele é uma característica que apresenta grande variação interpopulacional. Relativamente à sua previsão direta, vários genes de pigmentação foram apontados por FRUDAKIS (2008) como associados à variação do tom de pele em diversas populações, nomeadamente: MC1R, TYR, ASIP, MATP, SLC24A5, OCA2. O autor refere, no entanto, que a variabilidade observada nos referidos genes não é suficiente para explicar a variação fenotípica observada, pelo que, no caso do tom de pele, a inferência indireta é, a curto prazo, a melhor opção, visto, por um lado, ainda não estarem identificadas “todas as peças do puzzle” e, por outro, ser previsível a existência de diferentes mutações funcionais subjacentes à pigmentação da pele em diferentes populações (NORTON et al., 2007, STOKOWSKI et al., 2007) (assunto a desenvolver adiante).

Por seu turno, SPICHENOK et al. (2010) reportam a possibilidade de prever o tom de pele de um indivíduo acrescentando apenas um marcador (localizado no gene SLC24A5) à lista dos seis marcadores utilizados na previsão da cor dos olhos anteriormente descrito. Segundo os autores, a taxa de erro associada à previsão é muito baixa, mas, mais uma vez, a previsão por exclusão (não-branco e não-negro) e o facto de se a previsão não ter sido possível em 28% dos casos leva os autores a realçar a importância da adição de mais marcadores para melhorar os resultados.

Um dado interessante relativamente aos mais recentes conjuntos de marcadores descritos para a previsão direta de traços relacionados com a pigmentação é o facto de que alguns SNPs constam em mais do que um conjunto mesmo em relação a características diferentes: o rs12913832 (HERC2), o rs12203592 (IRF4) e o rs16891982 (SLC45A2) são comuns aos três conjuntos de marcadores e o rs 1800407 (OCA2) é comum a dois dos conjuntos (Tabela 10). Como será abordado adiante, este facto estará relacionado com a existência de variação fenotípica dos traços de pigmentação a nível interpopulacional.

Outro dado relevante da Tabela 10 diz respeito à localização dos SNPs em questão, sendo que, apesar de a maioria se localizar em regiões intrónicas ou apenas na proximidade dos referidos genes, alguns localizam-se em regiões codificantes interferindo com o aminoácido posteriormente incorporado na proteína e um deles (rs683) localiza-se na região 3'UTR, região que, apesar de não ser codificante, desempenha um importante

papel ao nível da regulação da tradução e estabilidade do mRNA. Os factos descritos levantam problemas ao nível da aplicação destes marcadores na prática forense devido a limitações legais, como será abordado adiante.

**Tabela 10 – Genes e respetivos SNPs presentes nos três mais recentes conjuntos de marcadores que permitem a previsão da cor dos olhos e cabelo e tom de pele.**

Gene (SNP) \ Referências	SPICHENOK et al. (2010)	LIU et al. (2009)	BRANICKI et al. (2011)
HERC2 (rs12913832)	O/P	O	C
IRF4 (rs12203592)	O/P	O	C
SLC45A2 (rs16891982)*	O/P	O	C
SLC45A2 (rs28777)			C
OCA2 (rs1545397)	O/P		
OCA2 (rs1800407)*		O	C
MC1R (rs885479)*	O/P		
MC1R_r			C
MC1R_R			C
ASIP (rs6119471)	O/P		
ASIP (rs2378249)			C
SLC24A5 (rs1426654)*	P		
SLC24A4 (rs2402130)			C
SLC24A4 (rs12896399)		O	
TYRP1 (rs683)**			C
TYR (rs1393350)		O	
TYR (rs1042602)*			C
EXOC2 (rs4959270)			C
KITLG (rs12821256)			C

\* – SNPs localizados em regiões codificantes que levam à alteração do aminoácido codificado; \*\* – SNP localizado na região 3'UTR – dados obtidos no site dbSNP (NCBI): <http://www.ncbi.nlm.nih.gov/projects/SNP/> consultado a 2-12-2011); O – previsão da cor dos olhos; P – previsão do tom de pele; C – previsão da cor do cabelo.

#### 3.1.2.4. ALTURA

A altura ou estatura do adulto é uma característica com influência de fatores hereditários estimada em 80% (VISSCHER et al., 2008) a 90% (HIRSCHHORN et al., 2009). Esta influência da componente genética foi inferida a partir de estudos familiares e também de estudos comparativos com gémeos monozigóticos e dizigóticos (sendo que os primeiros apresentam maior semelhança relativamente ao traço considerado) (HIRSCHHORN et al., 2009). Em 1886, Francis Galton demonstrou que a média de alturas dos pais está fortemente relacionada com a altura dos filhos, sendo um dado revelador no que toca à transmissão desta característica (AULCHENKO et al., 2009). Segundo AULCHENKO et al. (2009), este método de previsão da altura do indivíduo com base nas alturas dos pais explica cerca de 40% da variação da altura (ajustada ao sexo e idade), tendo demonstrado elevada precisão em termos discriminativos. No entanto, no âmbito forense estas informações não se encontram disponíveis pelo que os marcadores genéticos são a alternativa.

Esta característica fenotípica não está, porém, livre da influência ambiental, tendo-se verificado um aumento substancial da altura nos adultos nas últimas gerações que se deverá provavelmente a uma melhoria ao nível nutricional (HIRSCHHORN et al., 2009, LETTRE, 2009).

Apesar da consideravelmente significativa componente hereditária associada à estatura, os diversos estudos que têm vindo a ser realizados têm produzido resultados no mínimo dececionantes (Tabela 11). Exemplo disso é um estudo publicado em 2008 que identificou 20 marcadores que, em conjunto, explicam apenas 3% da variação observada numa amostra de 16.482 indivíduos com ancestralidade Europeia (WEEDON et al., 2008). Um artigo de revisão de LETTRE (2009) faz referência a um conjunto de 47 SNPs (descritos em estudos anteriores) que explicam cerca de 5% da variação da estatura observada em Caucasianos. Um estudo mais recente de LANGO ALLEN et al. (2010) recorreu a 180 loci envolvidos em vias metabólicas e defeitos do crescimento do esqueleto numa amostra de 50.074 indivíduos com ancestralidade Europeia, tendo conseguido explicar 10% da variação fenotípica. Em contraste com os estudos previamente referidos surge o trabalho de YANG et al. (2010) no qual é reportada a análise de 294.831 SNPs em 3.925 Europeus, sendo que os resultados do estudo apontam que 45% da variação fenotípica observada é explicada através da variação genotípica.



**Tabela 11 – Resumo dos dados relativos à percentagem da variação fenotípica associada ao número de marcadores utilizados de trabalhos relativos aos marcadores de DNA relacionados com a altura do adulto.**

Referência	Nº de Marcadores	% de variação fenotípica associada
Weedon et al. (2008)	20	3
Lettre (2009)	47	5
Lango Allen (2010)	180	10
Yang et al. (2010)	294.831	45

Os dados citados apontam para o facto de a estatura ser determinada por um número extremamente elevado de loci, cada um com fraco efeito, pelo que muitos estudos de associação não os detetam. Por outro lado, o poder de deteção dos estudos de associação é também afetado pela variância associada aos SNPs que, por sua vez é dependente do LD entre a variante causal e o SNP e a frequência e efeito da variante causal. Por outras palavras, variantes causais com fraco efeito ou alelos raros com forte efeito explicam uma pequena proporção da variação (YANG et al., 2010).

Os estudos anteriormente referidos são relativos a indivíduos de origem Europeia. No entanto, os dados obtidos nestes estudos poderão não ser reprodutíveis quando aplicados a amostras com origem ancestral diferente. O trabalho de LEI et al. (2009) refere isso mesmo (tendo como objecto de estudo a população Chinesa), tendo os autores concluído que a região estudada (região de aproximadamente 150kb que inclui dois genes: ZNP510 e ZNP782) se encontra “significativamente associada com a altura e a diferenciação interétnica em alguns loci pode contribuir para as observadas diferenças interétnicas relativamente à altura”.

Em suma, a estatura do adulto é uma característica fenotípica que apresenta uma forte hereditariedade. Os estudos realizados até à data ainda não permitiram identificar na totalidade as bases genéticas inerentes a esta característica, pelo que é esperado que nos próximos anos iniciativas como a criação do consórcio *Genetic Investigation of Anthropometric Traits* (GIANT) venham a revelar mais informações sobre este traço. No entanto, por se tratar de uma característica determinada pela ação de vários loci todos

eles com fraco efeito não se prevê num futuro próximo que venha a ser possível o desenvolvimento de kits para previsão da estatura com aplicação forense (KAYSER et al., 2009).

### 3.1.2.5. TRAÇOS FACIAIS

Segundo KOHN (1991), relativamente às medidas cefalométricas, verifica-se um elevado nível de hereditabilidade, o que sugere uma forte influência dos factores genéticos na determinação dos traços faciais.

O estudo recentemente publicado por BOEHRINGER et al. (2011) é descrito pelos autores como “uma primeira abordagem para a compreensão das bases genéticas da aparência facial humana”. O referido estudo estabelece a relação entre regiões do genoma previamente implicada numa patologia (fissura labial e/ ou palatina não-sindrómica, malformação popularmente conhecida como “lábio leporino” – Figura 36) e a variação de dois parâmetros: largura do nariz e largura da face (ou distância bizigomática). Recorrendo a 11 SNPs de 10 regiões genómicas diferentes analisados em duas amostras de dimensão e média de idade diferentes, com métodos de fenotipagem também diferentes, os autores obtiveram números bastante modestos em termos da percentagem da variação fenotípica que se pode explicar pela variação nos loci estudados (entre 0,57 e 2%), o que aponta, segundo os autores, para que os traços faciais sejam determinados por um grande número de genes. Outro aspeto interessante corresponde ao facto de que, nas amostras consideradas, os factores sexo e idade terem apresentado uma influência superior (entre 2 e 10%) à reportada relativamente aos factores genéticos analisados.



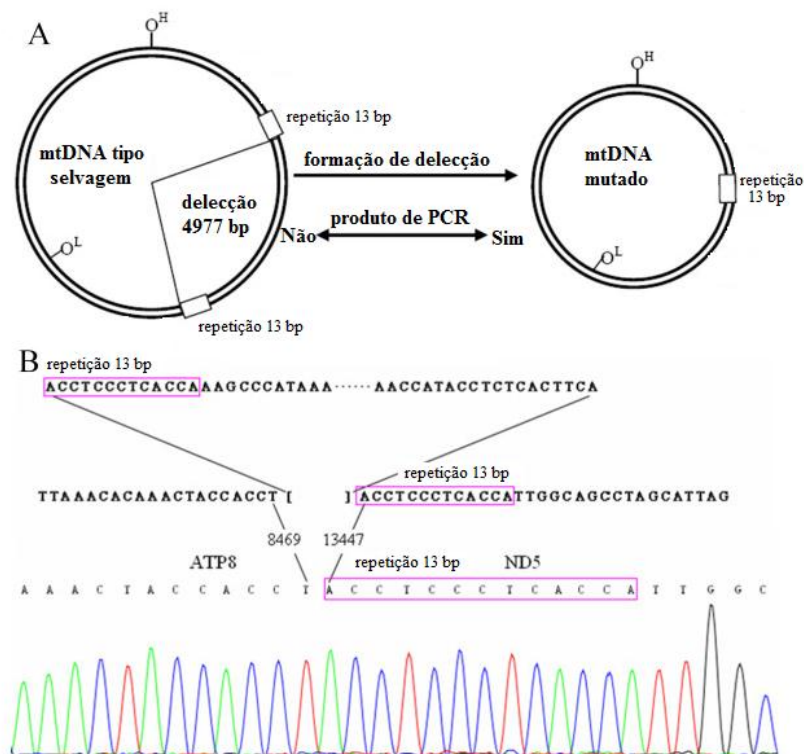
**Figura 36 – Criança com fissura labial ou “lábio leporino”. Fonte: <http://www.mundoeducacao.com.br/biologia/labio-leporino.htm> consultado a 02-12-2011)**

PULKER et al. (2007) referem, também, que algumas características faciais são tidas como monogénicas e constam na base de dados *Online Mendelian Inheritance in Man*® (OMIM): presença ou ausência de covinhas na face (OMIM 126100), presença ou ausência de covinha no queixo (OMIM 119000), linha frontal do cabelo com bico saliente na parte central da testa (pico da viúva) ou linear (OMIM 194000), presença ou ausência de pêlos nas orelhas (OMIM 139500), lóbulo da orelha preso ou solto (OMIM 128900).

### 3.1.2.6. IDADE

O envelhecimento é um processo natural que leva à ocorrência de alterações ao nível molecular nos órgãos e tecidos. Em termos de alterações verificadas ao nível do material genético podem salienta-se três fenómenos: a acumulação de deleções ao nível do mtDNA, a diminuição do tamanho dos telómeros (MEISSNER et al., 2010) e a diminuição do número de moléculas de DNA circular nas células T (ZUBAKOV et al., 2010).

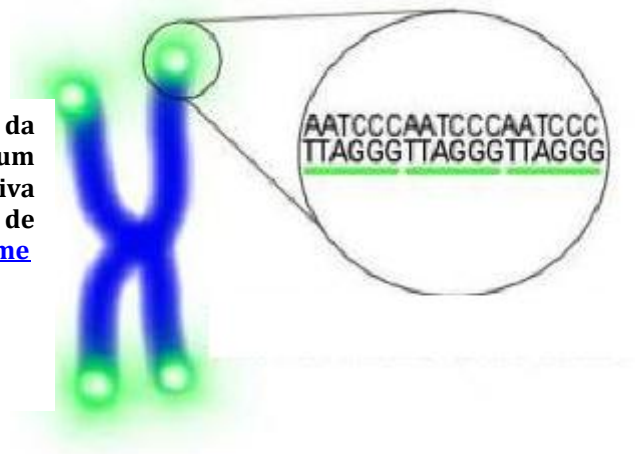
A acumulação de danos no DNA em particular a ocorrência de deleções ao nível do mtDNA (por se encontrar mais suscetível à ação nociva dos radicais livres libertados durante o processo da fosforilação oxidativa que tem lugar nas mitocôndrias) ocorre de forma aleatória em células únicas que, à medida que se replicam, promovem a expansão clonal dessas mesmas deleções. Por ter já sido alvo de extensa pesquisa em vários tecidos diferentes, o método de quantificação da deleção 4.977 bp ou “deleção comum” (Figura 37) é tido como método potencialmente útil na previsão da idade de um indivíduo. No entanto, este método apresenta algumas limitações importantes: observa-se uma grande heterogeneidade em termos de abundância deste tipo de deleção em diferentes tecidos e os resultados obtidos através da aplicação do método em amostras de sangue não são satisfatórios em termos de aplicação forense; é um método exigente, quer em termos técnicos, quer em termos de amostragem; ainda não está esclarecido o processo que leva à formação das deleções, ainda não se conhecem todas as variáveis com influência no processo e, por último, a margem de erro é de cerca de 40 anos, o que constitui a principal desvantagem do método (MEISSNER et al., 2010).



**Figura 37 - Detecção da deleção 4.977 bp do mtDNA. A - genoma mitocondrial humano sem e com a mutação. B - Perfil genético obtido através da sequenciação de produtos de PCR de indivíduos portadores da deleção. Adaptado de CHEN et al. (2011).**

Outro processo decorrente do envelhecimento é a diminuição do comprimento dos telômeros, que são estruturas constituídas por milhares de repetições em tandem (TTAGGG/CCCTAA)<sub>n</sub> associadas a proteínas, localizadas nas extremidades dos cromossomos como se encontra esquematizado na Figura 38. Trata-se de estruturas essenciais à sobrevivência e capacidade replicativa das células somáticas com capacidade de divisão. Ao longo do tempo, os telômeros sofrem uma diminuição em termos de comprimento como consequência do problema da finalização da replicação pela DNA polimerase, processamento dos telômeros a cada ciclo celular e exposição a stress oxidativo. Por outro lado as células possuem mecanismos de reparação que permitem adicionar sequências repetitivas na extremidade dos telômeros pela enzima telomerase, mecanismos esses que se vão tornando menos eficazes ao longo do tempo (MEISSNER et al., 2010). A avaliação do tamanho dos telômeros poderá, então, ser usada para prever a idade de um dador de uma amostra de DNA.

**Figura 38 – Esquema ilustrativo da localização dos telómeros num cromossoma e sequência repetitiva que os constitui. Adaptado de <http://diseasefreeto100.com/telomeres.htm> consultado a 25-10-2011.**



Este método apresenta, no entanto, várias desvantagens. Em primeiro lugar, a maioria das técnicas aplicadas na avaliação do comprimento dos telómeros são muito trabalhosas e bastante exigentes em termos de equipamentos para além de exigirem amostras de boa qualidade e elevada quantidade de DNA (southern blott, hibridização in situ com fluorescência quantitativa (Q-FISH) e de fluxo (Flow-FISH)) (MEISSNER et al., 2010). Apenas um método que recorre à PCR quantitativa (Q-PCR) permite a análise de amostras com reduzida quantidade de material; no entanto a reprodutibilidade e precisão deste método não é descrita como fiável (AUBERT et al., 2011, MEISSNER et al., 2010). Outro fator importante a considerar corresponde ao facto de que o estado de degradação da amostra influencia os resultados obtidos, ou seja, foram obtidos diferentes resultados em amostras do mesmo indivíduo com estados de degradação diferentes. Também o próprio processo de diminuição do comprimento dos telómeros não é homogéneo quer a nível tecidual quer a nível dos cromossomas dentro da mesma célula, pelo que se observam diferenças intraindividuais marcadas. Consequentemente, tal como no caso da quantificação da deleção 4.799 bp do mtDNA, a margem de erro deste método é muito grande, sendo que foram observados comprimentos de telómeros semelhantes em indivíduos cujas idades diferiam em 40-50 anos (MEISSNER et al., 2010).

O último fenómeno referido de alteração genética decorrente do envelhecimento (aumento do número de moléculas de DNA circular nas células T) tem por base o rearranjo de genes que codificam os recetores dos linfócitos T (*T-cell receptors* – TCR) como consequência da diminuição do tamanho do timo ao longo da vida de um indivíduo (JOBILING, 2011, ZUBAKOV et al., 2010). No sentido de gerar um conjunto abrangente de moléculas recetoras, cada linfócito T imaturo é sujeito a um rearranjo somático particular

no loci TCR durante o seu desenvolvimento que ocorre no timo. O processo de desenvolvimento envolve a deleção e circularização de segmentos de DNA do referido loci em moléculas de DNA episomal – os círculos de excisão de TCR de sinal repartido (sjTRECs). O número de sjTRECs presente nas células T varia de forma inversamente proporcional à idade, refletindo o processo de involução do timo, pelo que a quantificação dessas mesmas moléculas por Q-PCR permite prever a idade de um dador de uma amostra que contenha essas mesmas células (como por exemplo sangue, mas não é aplicável em amostras de saliva ou sémen) (ZUBAKOV et al., 2010). Em termos de vantagens, este método apresenta bons resultados em amostras com algum grau de degradação e não requer grandes quantidades de DNA (os autores referem que a quantidade mínima necessária é 5 ng). Por outro lado, existem ainda vários fatores (sexo, estado de saúde do sistema imunitário (incluindo doenças infecciosas (SIDA) ou leucemia) e ancestralidade biogeográfica) que poderão ter alguma influência nos resultados obtidos que necessitam ainda de ser investigados em pormenor. O desvio padrão deste método descrito por (ZUBAKOV et al., 2010) é de  $\pm 8,9$  anos o que, apesar de representar um intervalo menor do que os dos métodos anteriores é ainda considerável.

Em suma, foram apresentados três métodos de previsão da idade de um dador de uma amostra de DNA que têm por base três fenómenos decorrentes do envelhecimento. A principal limitação comum a todos os métodos reside na margem de erro da previsão, sendo que, apesar de tudo, o método de quantificação de sjTRECs é que apresenta um desvio padrão menor ( $\pm 8,9$  anos). Este método tem também a importante vantagem de não ser exigente relativamente à qualidade e quantidade da amostra e o facto de poder ser aplicado em amostras de sangue ao contrário dos primeiros dois métodos descritos que terão mais utilidade em casos de identificação de cadáveres por haver uma maior diversidade e disponibilidade de material.

## **3.2. Ancestralidade Biogeográfica**

A ancestralidade biogeográfica corresponde à associação entre a origem de um indivíduo e a localização geográfica de presumíveis antepassados. Não sendo possível testar populações ancestrais, as populações que atualmente vivem nesses locais são usadas como modelo (ROYAL et al., 2010).

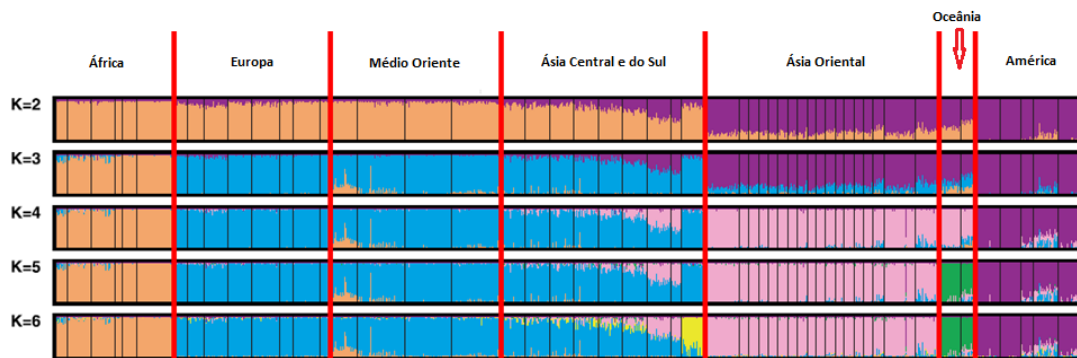
Como foi referido anteriormente, é previsível que a diversidade fenotípica interpopulacional se reflita a nível molecular. De facto, essas alterações ao nível molecular não vão envolver exclusivamente os loci diretamente envolvidos na característica em si, mas envolvem também toda a região próxima a esses mesmos loci. Neste sentido, (FRUDAKIS, 2008) refere que “a ancestralidade é usada para aceder por aproximação ao efeito de rede dos loci fenotipicamente funcionais com base na observação de que as características fenotípicas se encontram distribuídas em certa medida (quantificável) em função da ancestralidade”.

A inferência de características físicas a partir da Ancestralidade Biogeográfica é, então, uma forma de inferência indireta visto que a informação é obtida a partir de Marcadores Informativos de Ancestralidade (AIMs – *Ancestry Informative Markers*) e não a partir de genes diretamente responsáveis pelas características em estudo. Após a inferência da origem biogeográfica (através da análise dos AIMs) pode, então, ser aferida a relação entre a variação fenotípica e a origem biogeográfica relativamente a uma determinada característica (FRUDAKIS, 2008).

### **3.2.1. ANCESTRALIDADE INDIVIDUAL**

Como foi anteriormente referido, a variabilidade interpopulacional reflete-se, em termos genéticos, na variação de frequências alélicas, constituindo apenas uma pequena fração da variabilidade genética total. Em 2002, ROSERBERG et al. (2002) chegaram a uma classificação de ancestralidade biogeográfica constituída por cinco grupos cuja origem dos

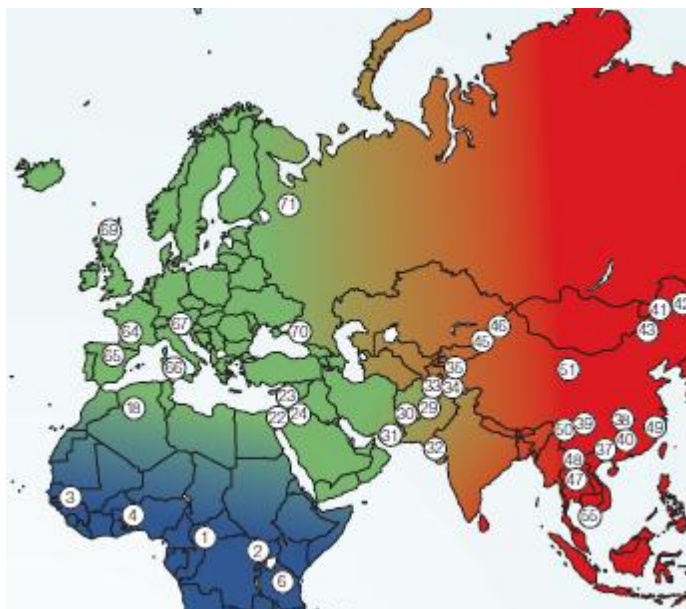
indivíduos corresponde a grandes regiões geográficas – África, Médio Oriente e Europa, Ásia, Oceânia e Américas (Figura 39). Esta classificação foi realizada com base na análise das frequências alélicas de 377 STRs autossómicos recorrendo ao programa estatístico STRUCTURE que, sucintamente, identifica subgrupos que apresentam frequências alélicas distintas, ou seja, classifica os indivíduos em K grupos, sendo que cada indivíduo pode ser membro de mais do que um grupo. Foi ainda identificado um sexto grupo (representado a amarelo na Figura 39) constituídos por indivíduos de uma população do norte do Paquistão (os Kalash). TISHKOFF et al. (2004) defendem que esta diferenciação se deve provavelmente a elevados níveis de endogamia e deriva genética que se terão feito sentir nesse grupo.



**Figura 39 – Estrutura populacional estimada recorrendo à análise estatística com o STRUCTURE dos dados obtidos através da aplicação de um painel de 377 STRs autossomais. Cada traço vertical do esquema representa um indivíduo; os traços vermelhos separam grupos de indivíduos com diferente origem geográfica. Adaptado de: ROSENBERG et al. (2002).**

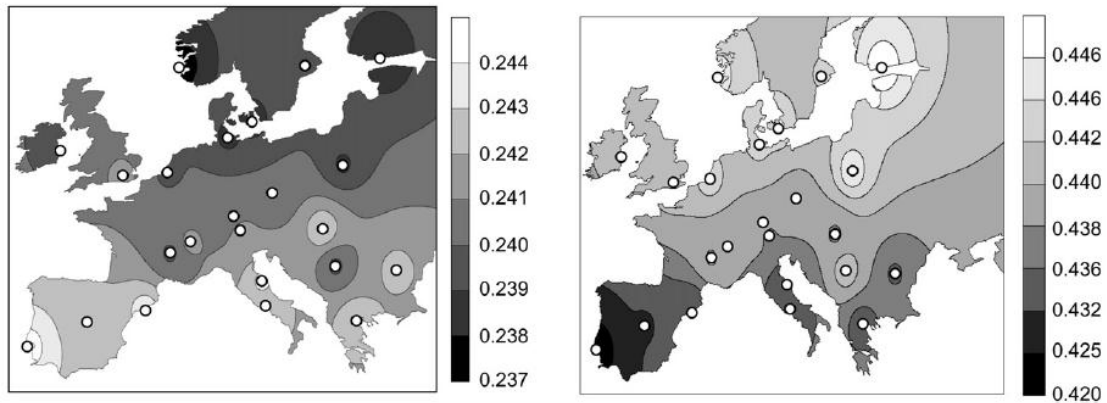
Apesar de a maioria dos indivíduos da mesma população partilharem o mesmo grupo major de origem há exceções, em especial nos grupos do Médio Oriente, Ásia Central e do Sul e Norte de África onde se verifica que determinadas populações são classificados em mais do que um grupo (BAMSHAD et al., 2004, KERSBERGEN et al., 2009, ROSENBERG et al., 2005, TISHKOFF et al., 2004). Este facto constitui um indicador de que a variação interpopulacional apresenta um gradiente contínuo, como se encontra representado na Figura 40.





**Figura 40 – Mapa representativo da variabilidade genética interpopulacional. Observa-se um gradiente contínuo de diversidade nos três continentes representados pelas três cores: Europa a verde, África a azul e Ásia a vermelho. Figura adaptada de BAMSHAD et al. (2004); números da figura representam algumas das populações estudadas no referido artigo.**

Este carácter contínuo que a variabilidade genética assume ao nível continental é também verificada ao nível intracontinental. Um estudo de LAO et al. (2008) concluiu que, apesar da escassa diferenciação genética observada ao nível das populações Europeias, verifica-se uma forte correlação entre as diferenças genéticas e a distância geográfica, para além de que o nível de heterozigotia é mais elevado nas populações do Sul do continente do que no Norte, sendo que este fator varia de forma inversamente proporcional ao LD (Figura 41). Segundo os autores, estes dados revelam que a variação genética na Europa assume um padrão gradual e contínuo em termos geográficos. Estas observações são consistentes com a história da população Europeia, em particular com a expansão populacional pré-histórica de sul para norte do continente e/ou um maior tamanho populacional efetivo na região sul relativamente ao norte da Europa.



**Figura 41** – Mapas representativos da distribuição geográfica do nível médio de heterozigotia (à esquerda) e LD (à direita) em 23 sub-populações Europeias. Fonte: LAO et al. (2008).

Os dados que têm vindo a ser descritos foram obtidos através da aplicação de painéis de AIMS. Assim, segue-se a descrição das principais características desses mesmos painéis.

### *3.2.2. PAINÉIS DE AIMS*

Ao longo da última década, vários estudos têm vindo a descrever diferentes painéis de marcadores desenhados com o propósito de classificarem geneticamente os indivíduos de acordo com a sua origem biogeográfica. Relativamente a esses KIDD et al. (2011) sugere a distinção de diferentes níveis pelo que, relativamente a esta matéria, os painéis podem ser classificados em três grupos (Tabela 12): identificação da origem ancestral continental, identificação da estratificação populacional dentro de determinadas áreas geográficas como a Europa e Ásia e a estimativa da proporção de mistura entre populações ancestrais relativamente a uma dada população ou grupo, como por exemplo a contribuição Africana e Europeia em Afro-Americanos e as contribuições Europeia, Nativo-Americana e Africana em populações Latinas e/ou Hispânicas.

**Tabela 12 – Lista de referências bibliográficas correspondentes a publicações de painéis de marcadores que permitem a determinação da origem biogeográfica de acordo com diferentes níveis de distinção.**

Níveis de distinção		Referências
<b>Origem ancestral continental</b>		ALLOCCO et al., 2007, ARMENGOL et al., 2009, BAMSHAD et al., 2003, FRUDAKIS, 2008, HALDER et al., 2008, KERSBERGEN et al., 2009, KOSOY et al., 2009, LAO et al., 2006, LI et al., 2008, LONDIN et al., 2010, NASSIR et al., 2009, PHILLIPS et al., 2007, RAAUM et al., 2010, ROSENBERG et al., 2005, ROSERBERG et al., 2002, SHRIVER et al., 2005, WATKINS et al., 2003
<b>Estratificação populacional dentro de determinadas áreas geográficas</b>	<b>Europa</b>	BAUCHET et al., 2007, LAO et al., 2008, NELIS et al., 2009, NOVEMBRE et al., 2008, SELDING et al., 2006, TIAN et al., 2009, TIAN et al., 2008b
	<b>Ásia</b>	KIM et al., 2005, THE HUGO PAN-ASIAN SNP CONSORTIUM, 2009
<b>Proporção de mistura entre populações ancestrais relativamente a uma dada população ou grupo</b>	<b>Proporção de ancestralidade Africana e Europeia em Afro-Americanos</b>	TANDON et al., 2011, YAEGER et al., 2008
	<b>Proporção de ancestralidade Europeia, Nativo-Americana e Africana em populações Latinas e/ou Hispânicas</b>	BONILLA et al., 2004a, HALDER et al., 2009, LEE et al., 2010, MAO et al., 2007, PENA et al., 2009, PIMENTA et al., 2006, PRICE et al., 2007, RIBEIRO-RODRIGUES et al., 2009, SANTOS et al., 2010, SELDIN et al., 2007, SILVA et al., 2010b, TIAN et al., 2007
	<b>Proporção de ancestralidade Europeia e Nativo-Americana em México-Americanos</b>	COLLINS-SCHRAMM et al., 2004

Os painéis anteriormente referidos apresentam características muito distintas nomeadamente no que toca ao número e tipo de marcadores, sendo que este pode variar entre 10 (LAO et al., 2006) e 650.000 (LI et al., 2008) e os marcadores estudados, apesar de, na maioria dos estudos, serem SNPs autossómicos, alguns trabalhos incluíram também marcadores do cromossoma X (KERSBERGEN et al., 2009, LI et al., 2008, SELDING et al., 2006, TANDON et al., 2011) e marcadores de linhagem (mtDNA e NRY) (BONILLA et al., 2004a, PENA et al., 2009). Outros ainda fizeram uso de STRs (LONDIN et al., 2010, PIMENTA et al., 2006, ROSENBERG et al., 2005, ROSERBERG et al., 2002) ou inserções Alu

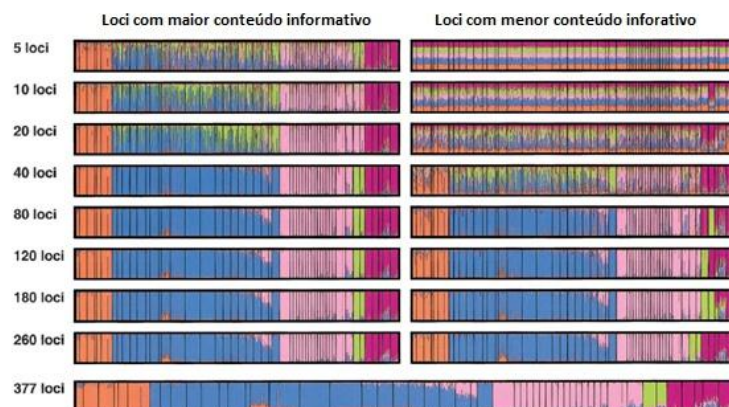
(RAAUM et al., 2010, WATKINS et al., 2003) ou ambos (BAMSHAD et al., 2003). As inserções Alu constituem a mais abundante classe de elementos curtos intercalados (*short interspersed elements* – SINEs) presente no genoma humano. Estes polimorfismos são constituídos por sequências diméricas com 300 bp que se propagam para novas localizações genômicas por retroposição (integração de sequências derivadas de RNA no genoma). Algumas inserções Alu estão relacionadas com doenças como a neurofibromatose tipo I, síndrome Apert e vários cânceres, mas a característica que desperta mais interesse no âmbito deste trabalho corresponde à variabilidade observada relativamente às diversas populações humanas, o que faz destes polimorfismos bons marcadores para estudos evolucionários e filogenéticos (WATKINS et al., 2003). Para além disso têm outras potenciais utilizações, nomeadamente ao nível da identificação de género e espécie (RAY et al., 2007).

A maioria dos artigos que constam na Tabela 12 foram realizados no âmbito clínico com um de dois propósitos: (1) evidenciar as diferenças genéticas relacionadas com a sub-estruturação da população e o seu efeito em termos de indução em erro aquando da realização de testes genéticos de associação, salientando a importância de uma avaliação objetiva da ancestralidade individual em vez da autoavaliação que é subjetiva (BAMSHAD et al., 2003, BAUCHET et al., 2007, KOSOY et al., 2009, LAO et al., 2008, NELIS et al., 2009, RAAUM et al., 2010, SELDING et al., 2006, TIAN et al., 2008a, TIAN et al., 2007, TIAN et al., 2009, YAEGER et al., 2008) ou (2) possibilidade de mapear características fenotípicas complexas associadas à ancestralidade biogeográfica, ou seja, conhecendo os loci nos quais há diferenciação em termos de frequências alélicas (onde se encontram identificados AIMS) poder-se-á chegar a loci (que se encontrarão próximos) responsáveis por determinadas características (como por exemplo doenças) nas quais há uma relação entre a característica e a ancestralidade do indivíduo (MAO et al., 2007, PRICE et al., 2007, SHRIVER et al., 2005, TANDON et al., 2011).

O mapeamento de características fenotípicas é também um ponto de interesse no campo forense uma vez que este pode ser o caminho para a identificação dos fatores genéticos que estão na base de características físicas complexas relativamente às quais se observa variação interpopulacional marcada (previsão direta). Enquanto esses fatores genéticos não forem totalmente (ou suficientemente) conhecidos, os dados relativos à ancestralidade biogeográfica podem ser usados para inferir indiretamente a provável aparência do dador de uma amostra de DNA (FRUDAKIS, 2008, KAYSER et al., 2009), o que, em determinados casos, poderia constituir uma causa provável para o pedido legal de

DNA de suspeitos que se encaixem no perfil, o que permitiria dar continuidade a investigações que não prosseguem por falta de suspeitos (FRUDAKIS et al., 2003). Para isso torna-se necessária a construção de painéis com características especiais, pois deverão possuir um número de marcadores reduzido (para que sejam práticos e, assim, utilizáveis na rotina) mas que, em simultâneo, assegurem uma elevada precisão em termos de classificação.

O processo de construção de um painel de marcadores é normalmente constituído por duas fases: inicialmente são seleccionadas amostras que representarão determinadas áreas geográficas; de seguida é aplicado um grande conjunto de marcadores (por exemplo um chip de DNA (GeneChip Human Mapping 500K Array Set (MAO et al., 2007), GeneMapping 10K Array Xba 131 (SHRIVER et al., 2005) ambos da Affymetrix Inc., Santa Clara, CA), vários chips de DNA (TANDON et al., 2011) ou conjuntos de marcadores que já haviam sido publicados (BAMSHAD et al., 2003, ROSERBERG et al., 2002, TIAN et al., 2007)) e é posteriormente realizada uma seleção de alguns marcadores. Essa seleção tem por base o conteúdo informativo de cada marcador que pode ser avaliado pelo  $I_n$  (informatividade para atribuição) ou outras medidas do conteúdo informativo de marcadores como o  $F_{ST}$ , a heterozigotia esperada ou a diferença absoluta entre frequências alélicas –  $\delta$ , entre outros (ROSENBERG et al., 2003), ou seja, são seleccionados os marcadores que apresentarem níveis elevados de diferenciação geográfica (em termos de frequências alélicas) nas amostras consideradas. Esta seleção permite obter melhores resultados com menor número de marcadores (Figura 42) do que se os marcadores fossem escolhidos aleatoriamente (ROSENBERG et al., 2003).



**Figura 42 – Representação esquemática da comparação da estrutura populacional de uma amostra em cinco grupos ( $k=5$ ) quando analisada com marcadores mais e menos informativos recorrendo ao STRUCTURE. Cada traço vertical representa o perfil de ancestralidade de um indivíduo que pode apresentar uma ou várias cores consoante o**

coeficiente de ancestralidade estimado. Os perfis de indivíduos de diferentes populações encontram-se separados por traços verticais pretos. A ordem pela qual os indivíduos se encontram representados é a mesma em todos os gráficos. Os gráficos presentes na parte superior da figura representam comparações entre os resultados obtidos com loci com maior informatividade vs loci com menor informatividade, sendo que o número de marcadores varia de linha para linha. O gráfico em baixo é o mesmo da Figura 1 de (ROSEBERG et al., 2002). Figura adaptada de ROSENBERG et al. (2003)

### 3.2.2.1. Painéis de AIMs com aplicação forense

Os painéis de marcadores a utilizar no contexto forense devem obedecer a uma série de critérios, em especial no que toca ao número de marcadores a analisar. Na prática, painéis com centenas ou milhares de marcadores não são exequíveis no âmbito de análises forenses devido a limitações técnicas e, principalmente, ao importante factor tempo.

Em 2003 a DNA Print Genomics lançou kits nos mercados forense (DNAWitness™ 2.5) e consumidores comuns (Ancestry by DNA 2.5) com o propósito de permitir inferir a ancestralidade genética de um indivíduo a partir do DNA (DNA PRINT GENOMICS). Trata-se de um painel (em ambos os kits) de 176 AIMs distribuídos pelos cromossomas que permite, então, determinar a qual ou quais dos grupos continentais (Africano sub-Sahara, Europeu, Este Asiático ou Nativo Americano) pertence o indivíduo, permitindo aceder à proporção das várias origens ancestrais em casos de mistura (Figura 43).

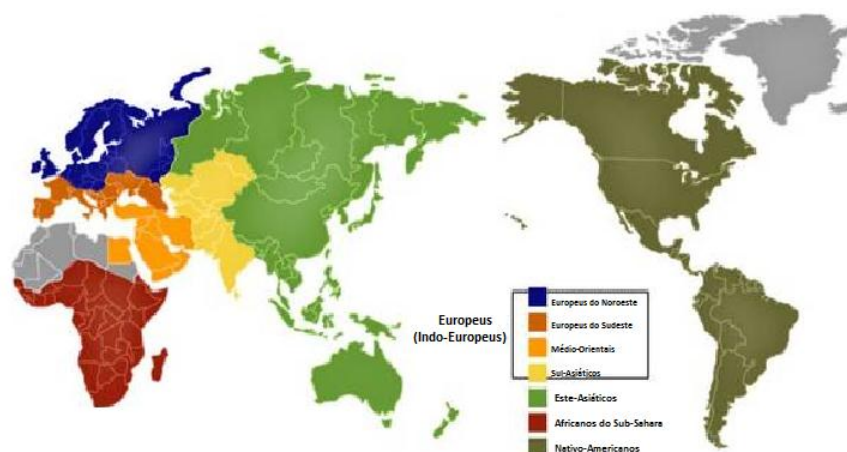
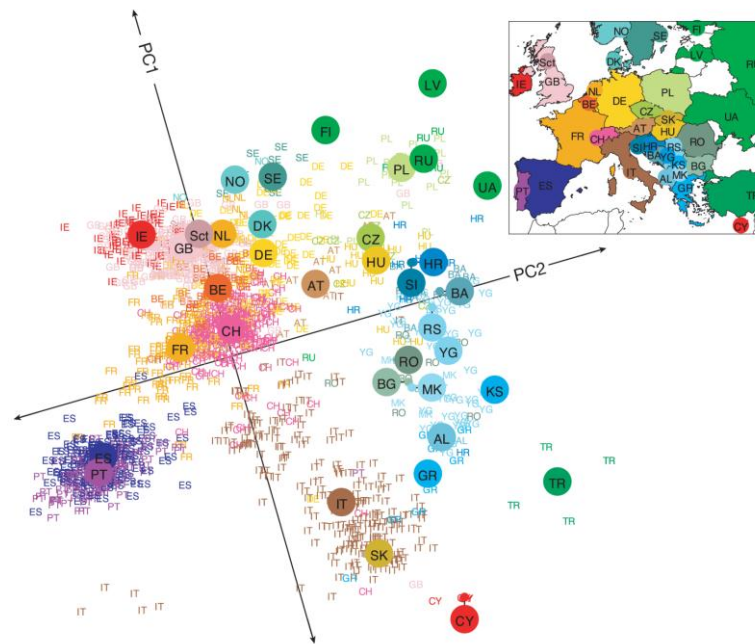


Figura 43 – Mapa representativo da caracterização da origem biogeográfica obtida com a aplicação dos kits Ancestry by DNA 2.5 e EurasianDNA 1.0. O primeiro distingue a origem biogeográfica continental sendo que as populações incluídas no conjunto Europeu ou Indo-Europeu são apenas distinguíveis recorrendo ao kit EurasianDNA 1.0. Figura adaptada do site: <http://www.dna-worldwide.com/ancestry-testing/ancestry-by-dna/> consultado a 19-04-2011.

Outro painel de inferência da ancestralidade biogeográfica desenvolvido no âmbito da análise de amostras forenses foi descrito por PHILLIPS et al. (2007) permitindo a distinção entre três grupos parentais (Africano, Europeu e Este Asiático) recorrendo a um conjunto de 34 AIMs.

Em termos de distinção intracontinental, a DNA Print Genomics desenvolveu também painéis de AIMs que permitem a aplicar em indivíduos com proporção ancestral Europeia superior a 50%. São então três painéis de marcadores, um deles lançado no mercado forense (EUROWitness™ 1.0) e os restantes dois disponíveis ao público geral (EurasianDNA 1.0 e EuropeanDNA 2.0). O EUROWitness™ 1.0 e o EurasianDNA 1.0 (Figura 43) fazem a distinção em quatro grupos: Europeus do Noroeste, Europeus do Sudeste, Médio Orientais e Sul Asiáticos. Já o EuropeanDNA 2.0 é um painel que tem por base o trabalho de (BAUCHET et al., 2007) e permite distinguir populações Europeias: Europeus do Sudeste (SEE – Arménios, Judeus, Gregos e Italianos), Ibéricos (IB – Portugueses e Espanhóis), Bascos (BAS – fronteira dos Pirinéus entre Espanha e França) e Europeus Continentais (CE – Bálticos, Polacos, Noruegueses, Finlandeses, Suecos e Russos). O que este teste ganha em sensibilidade e precisão perde em termos de aplicabilidade tanto sob o ponto de vista económico como prático visto ser composto por 1.349 marcadores (DNA PRINT GENOMICS).

Também o trabalho de NOVEMBRE et al. (2008) apresenta um painel de 197.146 AIMs aplicado a um total de 1.387 indivíduos oriundos de países Europeus (Figura 44). O estudo apresenta uma abordagem de atribuição baseada em regressão múltipla (PCA (*Principal Component Analysis*) em dois eixos) que permite determinar a origem biogeográfica dos indivíduos com uma impressionante precisão. Segundo os autores, esta abordagem permitiu posicionar 50% dos indivíduos num raio de 310 km relativamente ao seu local de origem, subindo a percentagem para 90% quando considerada uma distância máxima de 700 km (dados referentes a populações cuja representação foi superior a 6 indivíduos).



**Figura 44 – Representação esquemática dos dados genotípicos obtidos por (NOVEMBRE et al., 2008) recorrendo a uma abordagem de PCA em dois eixos. As pequenas etiquetas coloridas representam os indivíduos ao passo que os grandes círculos coloridos representam os valores médios relativamente a ambos os eixos para cada país. No mapa encontram-se representados os países e respetivas cores e etiquetas. AL – Albânia; AT, Áustria; BA, Bósnia-Herzegovina; BE, Bélgica; BG, Bulgária; CH, Suíça; CY, Chipre; CZ, República Checa; DE, Alemanha; DK, Dinamarca; ES, Espanha; FI, Finlândia; FR, França; GB, Reino Unido; GR, Grécia; HR, Croácia; HU, Hungria; IE, Irlanda; IT, Itália; KS, Kosovo; LV, Letónia; MK, Macedónia; NO, Noruega; NL, Holanda; PL, Polónia; PT, Portugal; RO, Roménia; RS, Sérvia e Montenegro; RU, Rússia; SCT, Escócia; SE, Suécia; SI, Eslovénia; SK, Eslováquia; TR, Turquia; UA, Ucrânia; YG, Jugoslávia. Fonte: NOVEMBRE et al. (2008).**

### ***3.2.3. Previsão de fenótipos a partir da Ancestralidade Biogeográfica – Método Indireto***

Sob o ponto de vista forense, o principal interesse do conhecimento da Ancestralidade Biogeográfica reside na possibilidade de, a partir deste, reconstruir a aparência física de um indivíduo, ou seja, este conhecimento permite fazer generalizações acerca de fenótipos que podem ajudar a identificar o dador de uma amostra de DNA. Esta possibilidade baseia-se no facto de algumas características apresentarem uma distribuição heterogénea entre as diversas populações (FRUDAKIS, 2008), sendo que essa variabilidade está diretamente relacionada com o processo de adaptação da espécie aos vários ambientes aos quais os seres humanos estiveram sujeitos ao longo do processo de



migração. Assim, à medida que as populações (ou fragmentos das populações) migravam de um ambiente para outro, as forças evolutivas foram moldando a composição genética dessas mesmas populações. Por outras palavras, as mudanças ambientais fizeram com que determinadas características conferissem maior adaptabilidade a quem as possuísse, o que, através do processo de seleção natural, fez com que com que esses indivíduos fossem alvo de seleção positiva, o que fez com que essas características se tornassem mais frequentes na população. Consequentemente essas alterações fizeram-se sentir a nível molecular onde se observam variações ao nível das frequências alélicas entre as diversas populações (LI et al., 2008).

São exemplos de características morfológicas com variabilidade fenotípica observada em função da localização geográfica da população a estatura, o índice de massa corporal, pilosidade excessiva, calvice, dobra epicântica dos olhos, arquitetura craniofacial, dimensões do crânio e dos dentes e as características associadas à pigmentação (cor dos olhos, cabelo e pele) (FRUDAKIS, 2008, KAYSER et al., 2009). Por seu turno, certas características são transversais a todos os seres humanos como o tamanho do cérebro, a inteligência e a capacidade de linguagem. Esta partilha deve-se, por um lado, ao facto de o *H. sapiens* ser uma espécie recente e, por outro, ao facto de essas características serem vantajosas em qualquer ambiente (TISHKOFF et al., 2004).

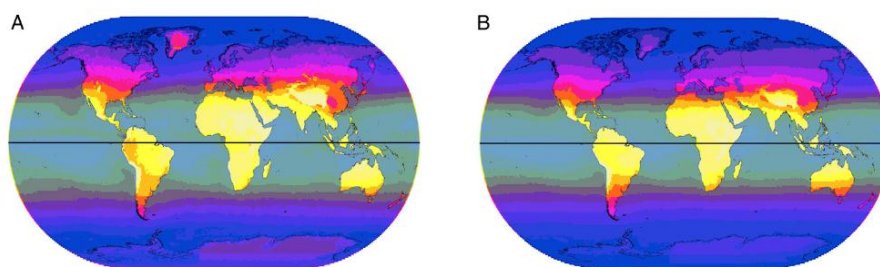
### 3.2.3.1. TRAÇOS RELACIONADOS COM A PIGMENTAÇÃO

O tom da pele (Figura 45) é, com foi já referido, um exemplo de característica que apresenta uma grande variabilidade interpopulacional (FRUDAKIS, 2008, JABLONSKI et al., 2000, 2010, STURM, 2009). Trata-se, segundo JABLONSKI et al. (2010) de um dos melhores exemplos de características humanas que se encontram sob a ação da seleção natural. Os autores referem também que o fenótipo observável ou nível de pigmentação constitutiva (pigmentação natural da pele sem a influência da radiação ultravioleta) varia gradualmente em função de dois fatores: (1) proteção da ação nociva da radiação ultravioleta que é mais intensa na região do equador e próxima ao equador (Figura 46), o que favorece, nessas regiões, o fenótipo de pigmentação constitutiva escura; (2) o fenómeno de produção de vitamina D através da fotossíntese induzida pela radiação UVB

favorece o fenótipo de pigmentação constitutiva clara nas regiões próxima aos pólos nas quais a intensidade dessa mesma radiação é menor (Figura 46). Nas latitudes intermédias, os níveis de UVB são sazonalmente elevados o que favorece o fenótipo de pigmentação constitutiva intermédia com capacidade de pigmentação facultativa (capacidade de desenvolver nível mais elevado de pigmentação da pela ação da radiação UV – bronzeado). Ainda segundo JABLONSKI et al. (2010), o nível de pigmentação constitutiva bem como a capacidade de pigmentação facultativa varia ao longo da vida do indivíduo, sendo que nas crianças se observa um nível de pigmentação menor do que o observado na fase final da adolescência/início da idade adulta (por volta dos 20 anos), altura em que os indivíduos atingem o nível máximo de pigmentação constitutiva bem como o nível máximo de capacidade de pigmentação facultativa em simultâneo com o seu período áureo de fertilidade. A partir da meia-idade em diante o nível de pigmentação constitutiva diminui gradualmente bem como o potencial de pigmentação facultativa devido à redução do número de melanócitos ativos. Existem também diferenças entre os sexos no que toca à pigmentação da pele, sendo que as mulheres têm normalmente um nível ligeiramente mais baixo de pigmentação constitutiva do que os homens devido, por um lado, à maior necessidade de vitamina D e, por outro, ao facto de em algumas culturas os homens demonstrarem preferência por mulheres com tom de pele mais claro – seleção sexual (JABLONSKI et al., 2000, 2010).

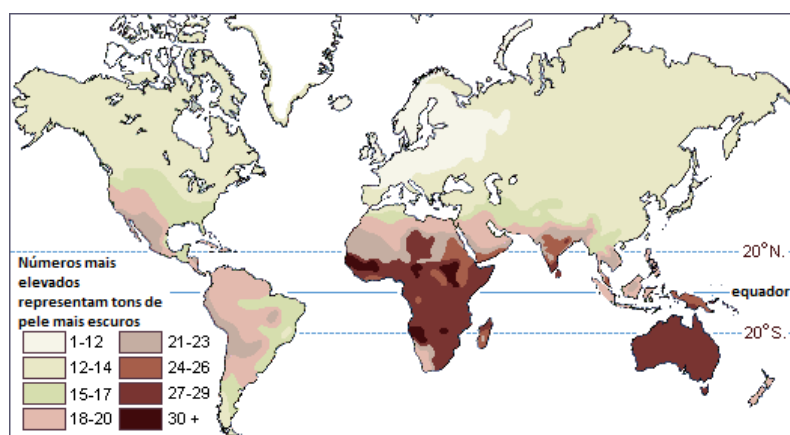


**Figura 45 – Fotografia de vários antebraços representando a diversidade fenotípica do tom de pele do mais escuro ao mais claro. Fonte: (STURM, 2009).**



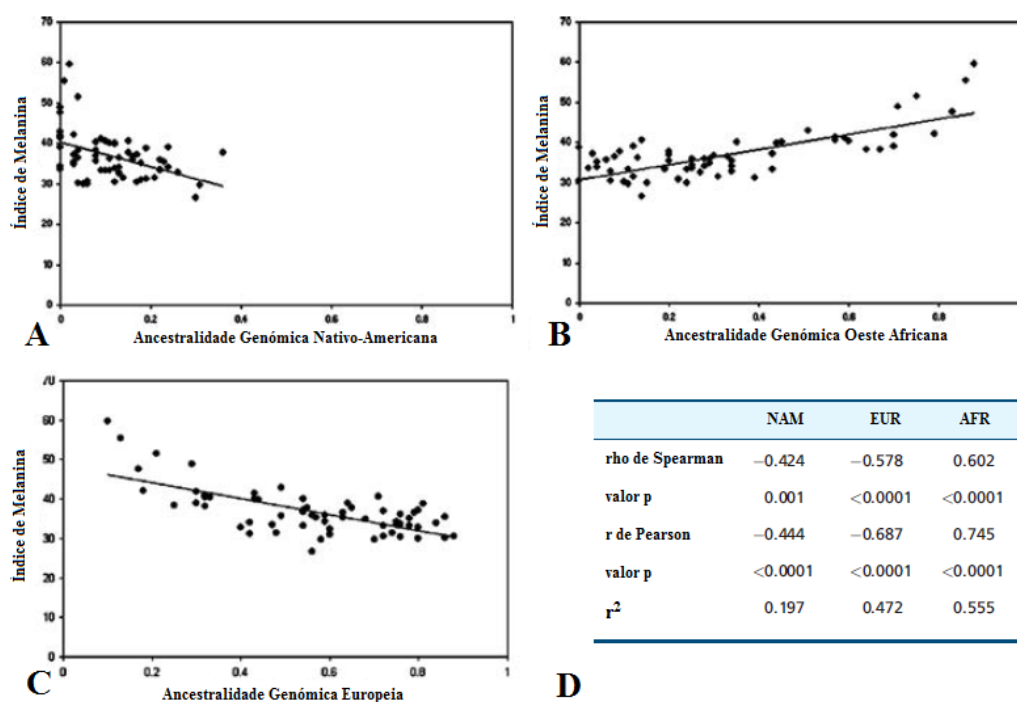
**Figura 46 – Níveis médios anuais de intensidade da radiação ultravioleta indicados por gradação de cor, sendo tanto mais elevados quanto mais clara a cor representada. A) Intensidade média anual de UVA (380nm). B) Intensidade média anual de UVB. Adaptado de JABLONSKI et al. (2010)**

O trabalho de LAO et al. (2007) no qual foram estudados 118 genes relacionados com a pigmentação da pele sugere que o tom de pele escuro terá, aparentemente uma origem única, refletindo o estado ancestral da espécie, ao passo que os tons de pele claros representam estados derivados com origem independente nas populações Europeia e Asiática. Na mesma linha, o trabalho de NORTON et al. (2007) refere que os genes ASIP e OCA2 desempenham um importante papel na determinação da pigmentação da pele a nível global, enquanto que os genes SLC24A5, SLC45A2 e TYR afetam apenas o fenótipo do tom de pele em indivíduos com ancestralidade Europeia. Assim, pressões ambientais semelhantes podem selecionar diferentes alelos em diferentes populações que se manifestam sob a forma de fenótipos semelhantes – evolução convergente (HANCOCK et al., 2011, NORTON et al., 2007, STOKOWSKI et al., 2007).



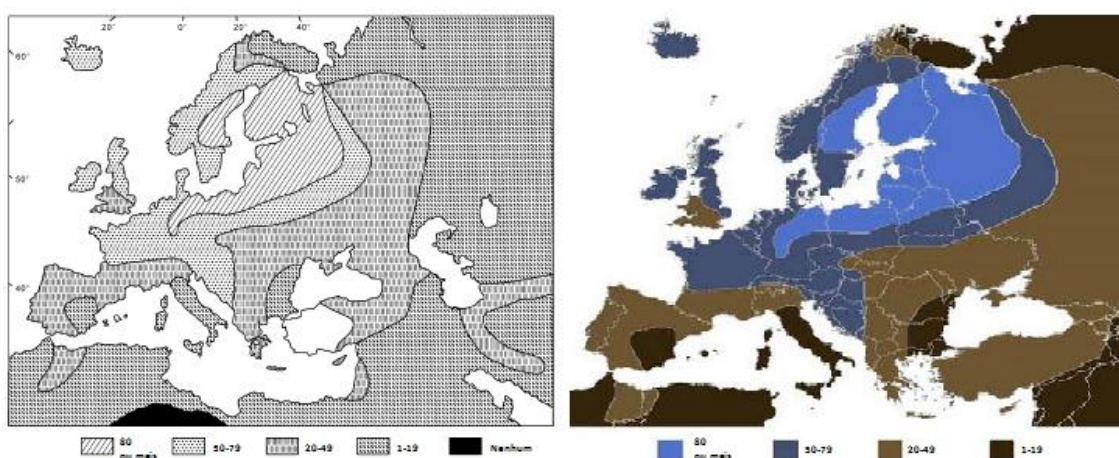
**Figura 47 – Distribuição do tom de pele nos seres humanos. Dados das populações nativas colhidos por R. Biasutti antes de 1940. Adaptado de FRUDAKIS (2008).**

FRUDAKIS (2008) refere que existe uma forte correlação entre os valores do índice de melanina (M) e a ancestralidade Africana (Figura 47), citando dados de dois trabalhos anteriores (BONILLA et al., 2004, PARRA et al., 2004). O autor alega também que os referidos trabalhos fazem uso de um reduzido número de marcadores, pelo que a estimativa de ancestralidade realizada a partir de um painel constituído por mais marcadores (171) permitiu, por um lado, aumentar a precisão da estimativa de mistura, e, por outro, aumentar a precisão da previsão do fenótipo (Figura 48). Essa previsão não está, porém, livre de erro, até porque a relação entre o valor do índice M e a ancestralidade não é exata ou rigorosa (Figura 48 – D) pelo que o autor adverte para a necessidade de haver um especial cuidado na forma como a informação é comunicada. O desenvolvimento de métodos de inferência direta à medida que se vão descobrindo as bases genéticas das características complexas permitirão ultrapassar esta desvantagem.



**Figura 48 – Correlação entre o Índice de Melanina e a Ancestralidade Genómica Nativo Americana (A), Oeste-Africana (B) e Europeia (C) determinada com base na aplicação de um painel de 171 AIMS descrito por (FRUDAKIS, 2008) a uma amostra de mulheres Porto Riquenhas. D – Dados estatísticos das correlações entre o Índice da Melanina e a Ancestralidade Genómica Nativo Americana (NAM), Oeste-Africana (AFR) e Europeia (EUR). Figura adaptada da obra referida. Figura adaptada de FRUDAKIS (2008).**

Relativamente à inferência indireta da cor dos olhos, FRUDAKIS, 2008 refere que existe associação entre essa característica e a ancestralidade continental, sendo que níveis de ancestralidade Africana a partir de 12%, Este Asiática a partir de 16% e Nativo-Americana superior a 25% se encontram associados a cores mais escuras da íris. Relativamente à ocorrência de olhos azuis, o autor refere que se encontra quase exclusivamente limitada a indivíduos com elevada proporção de ancestralidade Europeia (Figura 49) em particular Norte-Europeia (recorrendo ao painel de ancestralidade intracontinental EURO 1.0).



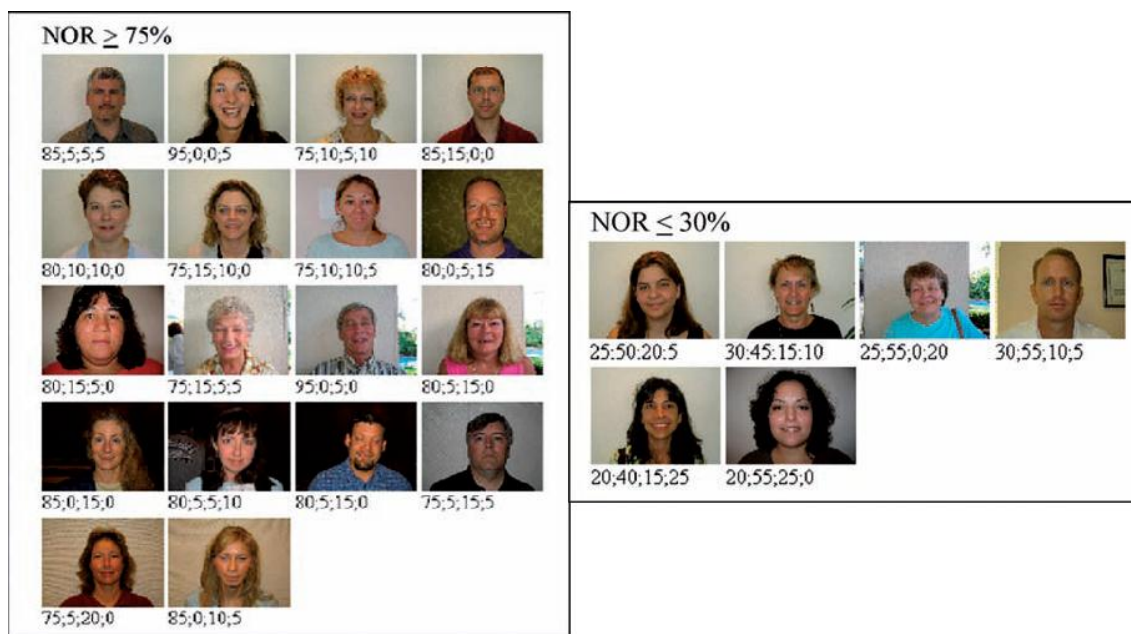
**Figura 49 – Distribuição geográfica das percentagens de frequência de olhos claros na Europa (e norte de África). Figurada esquerda adaptada de FROST (2006); figura da direita é sobreponível à da esquerda, representando um esquema de cores mais fácil de interpretar – adaptada do site [http://www.eupedia.com/europe/maps\\_of\\_europe.shtml#eye\\_colour](http://www.eupedia.com/europe/maps_of_europe.shtml#eye_colour), consultado a 13-04-2011.**

Os resultados obtidos através da utilização do painel EURO 1.0 descritos pelo autor que faz a distinção entre quatro grupos (Norte da Europa, Sul da Europa, Médio Oriente e Sul da Ásia) sugerem que a frequência de olhos claros decresce à medida que aumenta a distância geográfica relativamente a Norte da Europa, onde a frequência é mais elevada. Assim, o autor conclui que é expectável a possibilidade de prever a cor dos olhos de um indivíduo caso a proporção de ancestralidade Norte-Europeia ou Sul-Asiática seja elevada.

Em relação à previsão indireta da cor do cabelo, a única conclusão a que FRUDAKIS (2008) chegou, alertando, ainda assim, para a necessidade de realizar estudos recorrendo a conjuntos de dados mais extensos, foi que a cor natural do cabelo é aparentemente mais



clara em indivíduos cuja proporção de ancestralidade genómica Norte-Europeia é igual ou superior a 75% (Figura 50).



**Figura 50 – Fotografias de indivíduos que surgiram como resultado de uma pesquisa feita na base de dados EURO 1.0 correspondendo ao critério de pesquisa “NOR≥75%” (à esquerda) e “NOR≤ 30%” (NOR – ancestralidade genómica Norte-Europeia). Nesta pequena amostra de indivíduos, excluindo os que, por motivos relacionados com a idade, apresentam uma cor de cabelo alterada, é possível verificar que a ocorrência de cores de cabelo mais claras é mais frequente no grupo “NOR≥75%”. Adaptado de FRUDAKIS (2008).**

No sentido de averiguar quais os polimorfismos que estão na base da variação dos traços relacionados com a pigmentação (cor dos olhos, cor do cabelo e tom de pele) e se esses polimorfismos exibem frequências alélicas variáveis em diferentes populações, foram analisados os dados de 19 trabalhos nos quais foi estabelecida a relação entre genes e fenótipos (estudos de associação e painéis de previsão de cor dos olhos, cor do cabelo e tom de pele) ou entre genes de pigmentação e ancestralidade biogeográfica (Tabela 13). Os dados colhidos nesses mesmos trabalhos encontram-se sintetizados na Tabela A1, em anexo.

**Tabela 13 – Conjunto de 19 trabalhos relativos à pigmentação humana classificados de acordo com o tipo de estudo – previsão ou associação entre SNPs e fenótipo e relação entre SNPs de genes relacionados com a pigmentação e origem biogeográfica.**

Tipo de estudo		Referências
Relação entre SNPs e fenótipo	Associação	DUFFY et al., 2007, DUFFY et al., 2009, EIBERG et al., 2008, ERIKSSON et al., 2010, HAN et al., 2008, HANCOCK et al., 2011, KAYSER et al., 2008, LIU et al., 2010, MENGEL-FROM et al., 2010, STOKOWSKI et al., 2007, SULEM et al., 2008, SULEM et al., 2007, VALENZUELA et al., 2010
	Previsão	BRANICKI et al., 2011, SPICHENOK et al., 2010, WALSH et al., 2011*
Relação entre SNPs e origem biogeográfica		LAO et al., 2007, NORTON et al., 2007, STURM, 2009

**\*Os marcadores descritos por WALSH et al. (2011) são os mesmos que constam no trabalho de LIU et al. (2009).**

Nos referidos trabalhos consta um total de 96 SNPs em 26 genes. Identificou-se um conjunto de dez genes (MC1R, HERC2, IRF4, TYR, TYRP1, SLC24A4, SLC24A5, SLC45A2, OCA2 e KITLG) e sete SNPs (rs16891982 no SLC45A2, rs12913832 no HERC2; rs12203592 no IRF4; rs1800407 no OCA2; rs1286399 no SLC25A4; rs1042602 no TYR e rs1426654 no SLC24A5) citados em cinco ou mais artigos.

O SNP rs16891982 no SLC45A2 foi o SNP mais citado (13 referências) tendo sido reportado como relacionado com a variação fenotípica da cor dos olhos em sete trabalhos. Este polimorfismo é parte integrante de todos os painéis de previsão descritos e, tendo também sido reportada a ocorrência de variações em termos de frequências alélicas deste polimorfismo em diferentes populações em todos os estudos que estabelecem a relação entre genes de pigmentação e origem biogeográfica. Este marcador faz também parte do conjunto de marcadores descrito por PHILLIPS et al. (2007) para determinação da ancestralidade biogeográfica (SNPforID 34-plex) o que constitui outro indicador da diferenciação geográfica das frequências alélicas deste marcador.

O segundo SNP mais citado (11 trabalhos) foi o rs12913832 no HERC2, sendo que 8 das referências estabelecem a relação entre a variação alélica deste marcador e a variação da cor dos olhos. À semelhança do polimorfismo anteriormente referido, o rs12913832 integra todos os painéis de marcadores de previsão e é-lhe atribuído um papel preponderante na diferenciação entre olhos azuis e castanhos (MENGEL-FROM et al., 2010, WALSH et al., 2011). Apesar de não ser referido em nenhum dos trabalhos relativos

à relação entre genes associados a traços de pigmentação e ancestralidade biogeográfica, este SNP é também parte integrante do conjunto de marcadores SNPforID 34-plex.

Relativamente ao rs12203592 no IRF4, é reportado como relacionado com a cor dos olhos em 6 dos 8 trabalhos onde é referido, integrando três dos painéis de previsão incluídos na presente análise. Foram também observadas por STURM (2009) diferenças em termos de frequências alélicas deste polimorfismo em diferentes populações.

O rs1800407 no OCA2 é referido em 7 trabalhos, sendo descrita a associação entre este polimorfismo e a cor dos olhos em 4 dos artigos. Tal como foi referido relativamente ao SNP anterior, foram identificadas diferentes frequências alélicas deste marcador em diferentes populações por STURM (2009).

Com igual número de citações (7) surge o rs1286399 no SLC25A4, sendo que este SNP é descrito como influente na determinação da cor dos olhos e na cor do cabelo em 5 e 4 referências, respetivamente. Este marcador integra o sistema de previsão da cor dos olhos IrisPlex (WALSH et al., 2011) e, à semelhança dos anteriores, apresenta diferenças em termos de frequências alélicas deste polimorfismo em diferentes populações (STURM, 2009).

Para finalizar a análise dos polimorfismos mais citados, o rs1042602 no TYR e rs1426654 no SLC24A5 são apontados como genes influentes na determinação do tom de pele, tendo sido identificadas diferentes frequências alélicas em diferentes populações em ambos os casos (NORTON et al., 2007, STURM, 2009). O rs1426654 no SLC24A5 integra o SNPforID 34-plex (PHILLIPS et al., 2007).

### 3.2.3.2. MORFOLOGIA DO CABELO

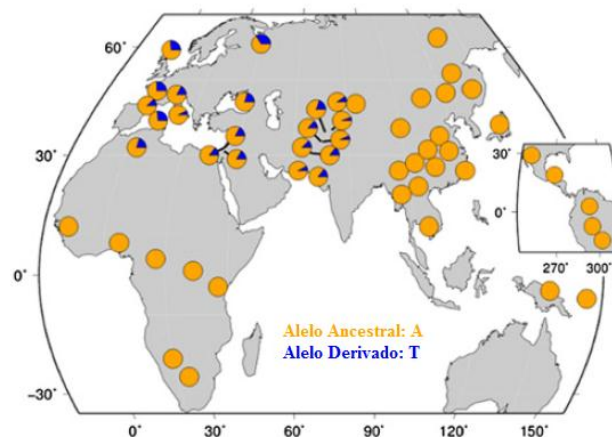
A morfologia do cabelo é determinada, por um lado, pela forma do folículo capilar (sendo que folículos retos originam fios de cabelo liso enquanto que folículos com dobra ao nível da base semelhante a um taco de golfe originam fios de cabelo encaracolado) e, por outro, pela assimetria em termos de espessura da bainha interna da raiz do cabelo cuja função é envolver e proteger o cabelo em crescimento na base do folículo (ERIKSSON et al., 2010).



Trata-se de uma característica com elevado grau de hereditariedade (MEDLAND et al., 2009) e variação interpopulacional substancial (WU et al., 2011). Também se observa variação intrapopulacional em particular na população Europeia na qual cerca de 45% dos indivíduos apresenta cabelo liso, 40% cabelo ondulado e 15% cabelo encaracolado (LOUSSOUARN et al., 2007).

Estudos relativos às bases genéticas da morfologia capilar em populações Asiáticas identificaram dois genes – EDAR (FUJIMOTO et al., 2008) e FGFR2 (FUJIMOTO et al., 2009) – que se encontram associados ao fenótipo de cabelo liso e espesso. Já em relação à população Europeia, um estudo de MEDLAND et al. (2009) conclui que a variação ao nível do gene TCHH (Figura 51) explica cerca de 6% da variação ao nível da morfologia capilar. Por seu turno, ERIKSSON et al. (2010) refere que regiões próximas dos genes LCE3E, WNT10A e OFCC1 se encontram relacionadas com a variação ao nível da morfologia capilar em Norte Europeus. Também o gene KRT71 foi apontado por SHIMOMURA et al. (2010) como estando relacionado com a característica em abordagem, sendo que os autores identificaram diferenças em termos de distribuições alélicas de sete SNPs deste mesmo gene relativamente às populações Europeia e Africana.

Assim, à semelhança do que foi referido em relação ao tom de pele, a evolução da morfologia capilar nas populações Europeia e Asiática ocorreu também de forma independente (MEDLAND et al., 2009).



**Figura 51 – Frequências dos alelos A e T do SNP rs 11803731 do gene TCHH com base no Human Genome Diversity Project. Adaptado de MEDLAND et al. (2009).**

### 3.2.3.3. ÍNDICE DE MASSA CORPORAL

Também o Índice de Massa Corporal apresenta variação em função da ancestralidade. FRUDAKIS (2008) refere que, segundo a regra de Bergman, esta característica é diretamente proporcional à distância ao equador uma vez que nos climas frios torna-se vantajoso possuir um rácio de volume/superfície corporal elevado por permitir uma menor perda de calor. Exemplos ilustrativos desta regra (Figura 52) são as populações da Etiópia e Quênia nas quais se observam indivíduos altos e magros comparativamente às populações do Ártico (os Inuit).



**Figura 52 – Exemplos ilustrativos da regra de Bergman. A – Família Erbore (Etiópia), fonte: <http://www.flickrriver.com/photos/41386913@N00/2240726561/> consultado a 04-10-2011. B – Crianças Inuits, fonte: <http://corpobrasil.blogspot.com/2010/05/o-povo-inuit.html> consultado a 04-10-2011 consultado a 04-10-2011.**

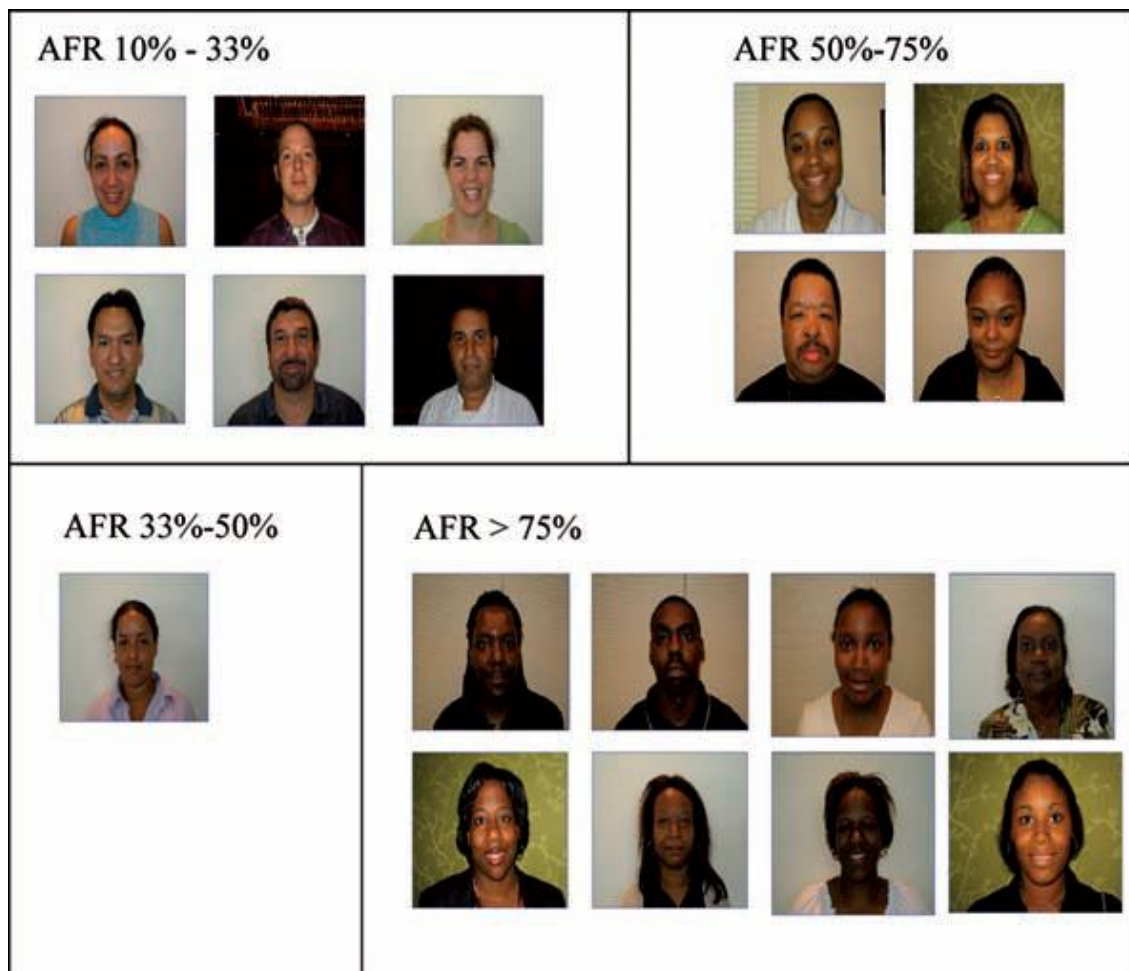
Associada à regra de Bergman surge uma outra – regra de Allen – que postula que membros mais curtos constituem uma característica vantajosa em climas frios sob o ponto de vista adaptativo por também maximizar a relação volume/superfície corporal. (FRUDAKIS, 2008) enuncia, no entanto, que existem exceções a estas regras (como por exemplo as populações de pigmeus Africanos) que sugerem que existem outros fatores em jogo, como por exemplo a seleção sexual, a adaptação a microclimas e até o tempo decorrido desde a fixação da população num determinado local. Estas exceções ilustram a necessidade de obtenção de dados concretos e de uma análise cuidada dos mesmos antes de proferir conclusões generalizadas.

#### 3.2.3.4. TRAÇOS FACIAIS

Por seu turno as características faciais constituem o conjunto de traços relativamente aos quais se observam as maiores diferenças entre populações, sendo que SMITH (2011) reporta a existência de uma correlação significativa entre o grau de semelhanças observadas relativamente à morfologia facial, do osso temporal e mandíbula em diferentes grupos e a distância geográfica. Este facto em conjunto com as observações obtidas a partir de estudos familiares que apontam para a existência de uma elevada influência de fatores genéticos na expressão destas mesmas características (PULKER et al., 2007 e VAN DAAL et al., 2007) leva a concluir, segundo FRUDAKIS (2008) que estes traços fenotípicos evoluíram, em certa medida, em função da filogeografia.

#### 3.2.4. FERRAMENTAS DE PREVISÃO INDIRETA DE FENÓTIPOS

O principal objetivo do método indireto de previsão de fenótipos é, como foi anteriormente referido, reconstruir a aparência física de um indivíduo a partir do ser perfil de ancestralidade. Para isso torna-se necessário compreender até que ponto uma determinada característica varia em função de fatores genéticos associados à ancestralidade, sendo que a melhor forma de o fazer é, segundo FRUDAKIS (2008), através da construção de uma base de dados. De facto, a obra referida descreve a criação e desenvolvimento de uma base de dados na qual foram inseridos, dados relativos à caracterização genotípica e fenotípica (física, linguística, cultural, religiosa e política) de cada indivíduo, estando presentes na mesma indivíduos de várias populações. Esta abordagem permitiu, por um lado, identificar a existência ou não de associações estatisticamente significativas entre os diversos dados colhidos (por exemplo, estabelecer relações entre ancestralidade Africana e tom de pele ou perceber a partir de que percentagem de ancestralidade um indivíduo apresenta características fenotípicas associadas a esse grupo ancestral – Figura 53) e, por outro, aplicar esse conhecimento na investigação de casos concretos.



**Figura 53 – Utilização de base de dados para estabelecimento de relações entre diferentes proporções de origem biogeográfica Africana e características fenotípicas. Fonte: FRUDAKIS (2008).**

No âmbito da aplicação do conhecimento relativo à relação entre ancestralidade biogeográfica e fenótipos em casos forenses, FRUDAKIS (2008) descreve duas aplicações com particular interesse – reconhecimento facial reverso e descrição étnica ou geográfica reversa. A primeira aplicação consiste numa ferramenta que permite obter um conjunto de imagens ou uma imagem gerada por um software com base em fotografias presentes na base de dados de indivíduos cujo perfil se enquadra nos critérios de pesquisa (Figura 54). O reconhecimento facial reverso permite, assim, obter uma imagem da provável aparência do dador de uma amostra através do seu perfil de ancestralidade com base numa coleção de dados objetivos.

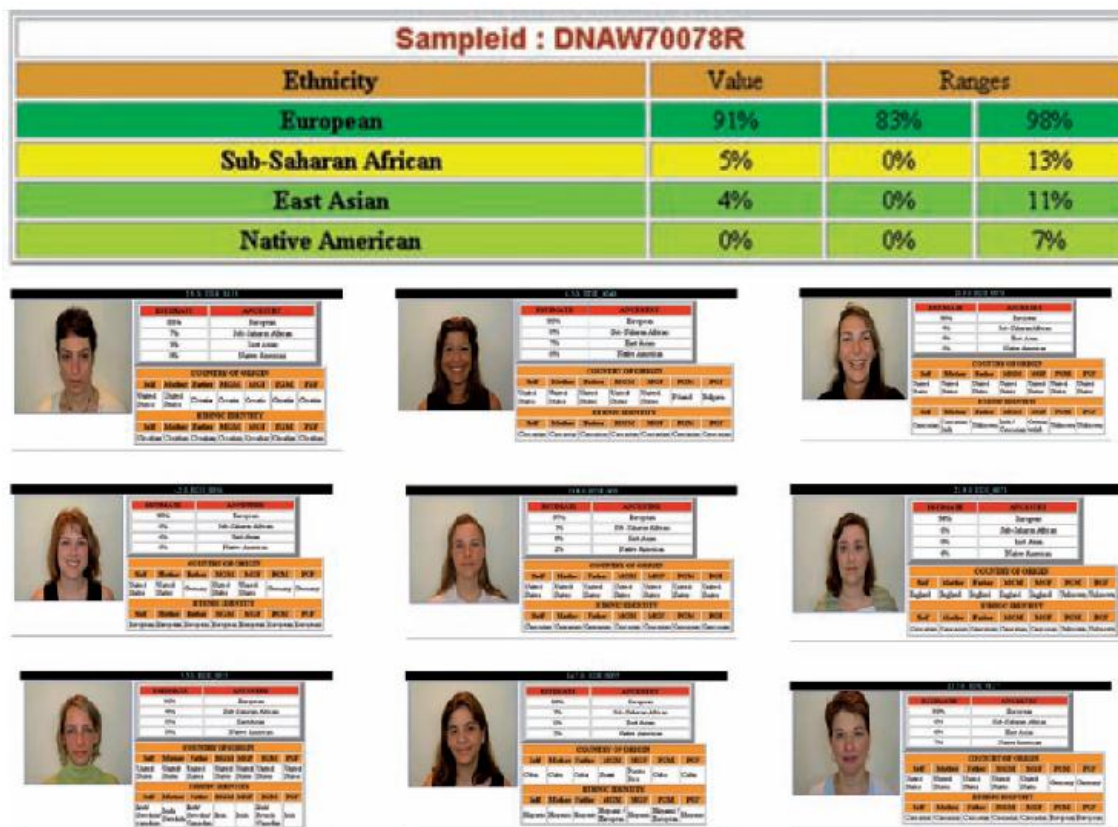


Figura 54 – Exemplo de pesquisa (em cima) e nove resultados obtidos (em baixo) na base de dados referida no trabalho de FRUDAKIS (2008) recorrendo a uma estimativa de mistura individual de ancestralidade genómica para um indivíduo Caucasiano do sexo feminino. Nesta base de dados constam, relativamente a cada indivíduo, para além dos valores estimados de mistura individual de ancestralidade genómica, valores relativos a diversos traços fenotípicos e dados demográficos (país de origem de familiares e respetivas descrições étnicas). A pesquisa fornece um conjunto de indivíduos cujos valores estimados de mistura individual de ancestralidade genómica correspondem aos intervalos estabelecidos para o dador desconhecido da amostra em análise. Adaptado de FRUDAKIS (2008).

Relativamente à descrição étnica ou geográfica reversa, constitui outra ferramenta na qual, à semelhança da anterior, é feita uma pesquisa na base de dados inserindo um determinado perfil, obtendo-se posteriormente dados relativos à autodescrição étnica ou racial, língua nativa, ou outros dados que constem nesses mesmos perfis. Assim, caso um número significativo de resultados seja consistente (os perfis apontem todos para a mesma etnia, a mesma língua nativa, a mesma religião, etc.), pode dizer-se que há uma grande probabilidade de que o dador da amostra que tencionamos identificar partilha dessas mesmas características. Obviamente, o valor dessa alegação será dependente da quantidade de dados na qual se baseia, pelo que quanto maior e mais representativa for a base de dados, mais precisos (e consequentemente úteis) serão os resultados.

O autor reconhece, no entanto, que o método indireto de previsão de fenótipos tem algumas desvantagens. As características que têm vindo a ser abordadas são traços complexos determinados não por um, mas por um conjunto de genes (traços poligénicos) que se encontram distribuídos pelo genoma. Assim, entre indivíduos com o mesmo perfil de mistura observar-se-á sempre variabilidade proveniente da complexidade genética subjacente ao traço (dominância, epistasia) e/ou erro de estimativa de proporção de mistura, sendo que essa variação não é uniforme quer sob o ponto de vista da característica fenotípica quer sob o ponto de vista do perfil de mistura. Essa variabilidade é, no entanto quantificável, o que permite perceber quando é que uma generalização (como por exemplo: “Indivíduos com proporção de mistura Oeste Africana apresentam um tom de pele mais escuro do que os indivíduos sem qualquer proporção de mistura Oeste Africana”) pode ou não ser feita com base em dados concretos.

### 3.3. CASOS PRÁTICOS

O primeiro caso forense descrito no qual foi aplicado o conhecimento relativo à previsão indireta de fenótipos ocorreu em 2002 na mesma altura em que foi desenvolvido pela DNAPrint Genomics um painel de 71 AIMs (DNAWitness™ 1.0). O referido painel foi, então, aplicado no âmbito da investigação de um caso de homicídio em série no Louisiana. O perfil de identificação do perpetrador dessas cinco violações/homicídios não encontrou correspondência na base de dados do FBI – *Federal Bureau of Investigation* (CODIS - *The Combined DNA Index System*) e a ansiedade e mau estar social a nível nacional, mas em particular na área geográfica na qual os crimes tinham tido lugar crescia exponencialmente. Duas testemunhas oculares alegavam ter visto um homem Caucasiano na vizinhança de um dos locais de crime. Um ano depois do início da investigação, a determinação do perfil de ancestralidade biogeográfica do indivíduo apontou outra direção; com uma proporção de ancestralidade Africana a sul do Sahara de 85% e 15% Nativo-Americana, foi apontada a probabilidade de que o tom de pele do suspeito seria mais escuro até do que a média dos indivíduos Afro-Americanos, o que levou a uma total mudança de rumo da investigação. Um mês depois da aplicação do teste de ancestralidade e após a reconsideração de suspeitos que haviam sido excluídos por não corresponderem



ao reportado pelas testemunhas, foi apanhado o perpetrador das violações/homicídios FRUDAKIS (2008).

Vários outros casos ocorridos nos Estados Unidos da América e até no Reino Unido são relatados por FRUDAKIS (2008), sendo que em comum têm o facto de a determinação do perfil de ancestralidade geográfica das amostras de local de crime ou até de corpos em avançado estado de decomposição, permitiu a poupança de tempo e outros recursos importantes para as investigações.

Em Espanha a determinação da origem biogeográfica de indivíduos alegadamente envolvidos nos atentados de 11 de Março de 2004 (Figura 55) recorrendo a marcadores da NRY e mtDNA combinados com um conjunto de 34 marcadores autossomais ajudou a restringir o número de suspeitos a considerar na investigação bem como a sua eventual filiação relativamente a duas organizações terroristas que poderiam estar envolvidas no atentado (PHILLIPS et al., 2009).



**Figura 55 - Atentado de 11 de Março de 2004 em Madrid. Fonte: <http://comunidad.laprovincia.es/galeria-multimedia/general/Galeria-fotos/600/5.html> consultado a 18-10-2011.**

### 3.4. A POPULAÇÃO PORTUGUESA

Ao longo dos anos têm sido publicados vários estudos relativos às frequências genéticas de vários marcadores observadas em populações Portuguesas. A Tabela A2 (em anexo) sistematiza a informação relativa ao tipo de marcadores estudados em 68 estudos desse género publicados a partir do ano 2000 e a Tabela A3 (em anexo) classifica esses mesmos estudos de acordo com as populações análise. Segue-se a descrição dos principais fatos relatados nos referidos estudos de acordo com a localização no genoma dos marcadores analisados.

#### 3.4.1. MARCADORES AUTOSSÓMICOS

Os trabalhos relativos a marcadores autossómicos incluídos nos kits de identificação forense (AmpF/STR® Profiler Plus™ – Applied Biosystems (adiante designado apenas por “Profiler”) – e/ou PowerPlex® 16 System – Promega (adiante designado apenas por “Powerplex”)) descrevem as frequências alélicas identificadas em populações Portuguesas, avaliando parâmetros como o poder de discriminação e exclusão e o nível de heterozigotia, que são sempre elevados em todas as populações avaliadas. Foram, porém, identificadas diferenças em termos de frequências alélicas em marcadores de ambos os kits. Em relação ao Profiler, AMORIM et al. (2001) afirma, com base em dados de uma amostra do Norte de Portugal, que o perfil genético dessa população é “tipicamente Caucasiano”, com diversidade genética inferior à reportada pela base de dados nacional de DNA dos Estados Unidos da América. Foram, no entanto, identificadas diferenças estatisticamente significativas ao nível dos alelos menos frequentes de 3 loci (D8S1179, D21S11 e D13S317) entre a população Portuguesa (amostra maioritariamente constituída por indivíduos do sul do país) e Italianos (GEADA et al., 2000). Em relação ao kit Powerplex, apesar de a população Portuguesa ser considerada homogénea (PINHEIRO et al., 2005), foram reportados alelos nas populações da Madeira e Açores que não foram observados no Continente (FERNANDES et al., 2002b, VELOSA et al., 2002).

Em relação a trabalhos nos quais foram estudadas as frequências alélicas de outros conjuntos de STRs (a maioria dos quais analisaram marcadores incluídos nesses mesmos



kits), foram identificadas diferenças entre populações: FERNANDES et al. (2002a) reportam diferenças em termos de frequências alélicas do locus TPO da população da Madeira relativamente às populações Centro e Sul de Portugal Continental; do locus VWA entre a Madeira e as populações Norte e Sul e entre os Açores e a população Sul; do locus CD4 entre a população Norte de Portugal e as restantes populações; os autores referem ainda que a população Centro de Portugal é a que apresenta menos diversidade. O trabalho de GUSMÃO et al. (2010) aponta para um nível de diversidade relativamente aos STRs autossómicos inferior na população Cigana Portuguesa do que o observado na população geral (76,7% vs 79%). O mesmo estudo refere que a redução da diversidade genética nos grupos Ciganos em relação à população envolvente aparentemente varia de acordo com um gradiente Este-Oeste, em paralelo com o seu padrão de dispersão pela Europa e que se observa uma elevada diferenciação genética entre grupos Ciganos, mais elevada do que as diferenças observadas entre cada grupo e respetiva população geral do mesmo país. A elevada heterogeneidade dever-se-á ao forte efeito da deriva genética e limitado fluxo genético intergrupo. Ainda segundo os autores, as elevadas taxas de componentes não-Indianos observadas nos pools genéticos das populações Ciganas indicam a ocorrência de intenso influxo genético das populações circundantes, o que contraria a ideia de que a comunidade Cigana é endógama.

LAGOA et al. (2008) reportam as frequências alélicas de um conjunto de mini-STRs na população do Norte de Portugal, sendo que, essas frequências se assemelham às observadas noutras populações Europeias (Espanhola, Italiana) e na população Caucasiana dos Estados Unidos da América, sendo, no entanto, diferentes das observadas em populações Afro-Americanas e Asiáticas.

Em relação à análise de marcadores da região HLA, observou-se uma grande variabilidade alélica e haplotípica ao nível da população Continental Portuguesa (Norte, Centro e Sul) (SPÍNOLA et al., 2002). O mais elevado nível de diversidade genética nesta região do genoma foi, no entanto, observado na população da Madeira (SPÍNOLA et al., 2002, SPÍNOLA et al., 2006).

No que toca a marcadores do tipo SNP noutras regiões que não o HLA, PEREIRA et al. (2008) descrevem as frequências alélicas de um conjunto de marcadores incluídos no SNPforID 52-plex nas populações Norte, Centro e Sul de Portugal como sendo muito semelhantes relativamente a praticamente todos os loci. Assim, os autores descrevem a população Portuguesa como homogénea e enfatizam a utilidade deste kit no âmbito da

identificação forense devido ao seu elevado poder de discriminação. COSTA et al. (2008), por seu turno enfatizam o elevado poder de discriminação de um sub-conjunto de apenas 10 SNPs dos 52 incluídos no SNPforID 52-plex e a sua grande utilidade como complemento aos marcadores STR em casos complexo de determinação de paternidade.

O estudo de BRANCO et al. (2006) relativo às frequências alélicas de 6 inserções Alu na população Açoriana e Portuguesa Continental conclui que a população Portuguesa é muito semelhante a outras populações Europeias, sendo que os dados apontam para uma grande proximidade entre Açorianos, população Continental, Catalães, Andaluzes, Marroquinos e Algerianos.

### *3.4.2. MARCADORES DO CROMOSSOMA X*

Em relação ao estudo de frequências alélicas de marcadores STR do cromossoma X, vários estudos (CARVALHO et al., 2011, GUSMÃO et al., 2009, PEREIRA et al., 2007) referem um elevado nível de homogeneidade genética no que toca à população Portuguesa, não se verificando diferenças significativas a nível regional ou relativamente a outros países, pelo que o conjunto de 10 STRs do cromossoma X visado nesses estudos é tipo como ferramenta útil na prática forense, em particular, em casos de determinação de parentesco. Por seu turno, SILVA et al. (2010) reportam a ocorrência de frequências alélicas significativamente diferentes entre os grupos Ocidental e Oriental do arquipélago dos Açores o que é indicativo da existência de heterogeneidade genética da população Açoriana, pelo que os autores defendem que é pertinente a criação de uma base de dados específica para esta população relativamente aos marcadores em análise. Foi também reportado por VALENTE et al. (2009) a ocorrência de um nível de heterozigotia reduzido na população Cigana Portuguesa comparativamente à população nacional geral (0,721 vs 0,745) que, segundo os autores, dever-se-á a um menor tamanho populacional efetivo, endogamia e forte efeito da deriva genética característicos dessa mesma população.

Já em relação a marcadores do tipo SNP, PEREIRA et al. (2011) classificam a população Portuguesa como geneticamente homogénea, não se verificando diferenças significativas a nível regional ou relativamente a outros países, em particular países do Mediterrâneo.

### **3.4.3. MARCADORES DO CROMOSSOMA Y**

No que diz respeito aos marcadores STR do cromossoma Y, os estudos relativos às frequências alélicas do conjunto de 17 marcadores do kit de amplificação AmpF/STR® Yfiler™ (Applied Biosystems) (ALVES et al., 2007, BENTO et al., 2008, BENTO et al., 2009b, GUSMÃO et al., 2008b, PONTES et al., 2007) e do kit Y-PLEX™ (ReliaGene) (FERNANDES et al., 2003), revelam que a população Portuguesa apresenta uma grande diversidade de haplótipos, tendo sido observadas diferenças a nível regional, com os Açores a apresentarem frequências alélicas significativamente diferentes das observadas nas populações Norte e Madeira (FERNANDES et al., 2003). Foram também reportadas diferenças entre a população Cigana portuguesa e a restante população do país (GUSMÃO et al., 2008b). Em termos de comparação com países próximos, PONTES et al. (2007) referem a existência de haplótipos comuns entre a população Norte de Portugal e Norte de Espanha, ao contrário do que se verifica entre as populações Norte de Portugal e Noroeste de Itália no que toca aos loci analisados.

Em relação a outros conjuntos de marcadores STR do cromossoma Y, GONZÁLEZ-NEIRA et al. (2000) referem, com base em dados de várias populações da Península Ibérica, que existe diferenciação populacional significativa entre a população de Valença e a população do Norte de Portugal. Por seu turno, BELEZA et al. (2006) enfatizam o elevado nível de diversidade genética observado na população do Alentejo comparativamente às restantes regiões do território nacional, que, segundo os autores, dever-se-á à abundância de um elevado número de linhagens diferentes de ancestralidade recente nessa mesma região. Foi também reportada a ocorrência de um haplótipo típico do Norte da Europa na população da Madeira que não foi observado na Península Ibérica por FERNANDES et al. (2001). Já GUSMÃO et al. (2008b), descreve a diversidade haplotípica observada na população Cigana Portuguesa como baixa (94,37%) comparativamente ao observado nas populações Europeias (99,9%). Refere ainda que as populações Ciganas de Portugal, Espanha, Lituânia e Bulgária são mais semelhantes entre si do que relativamente às populações gerais dos respetivos países, sendo que a proximidade genética observada entre a população Cigana Portuguesa e as restantes é proporcional à distância geográfica.

No que toca aos estudos da variação genética do cromossoma Y recorrendo à análise de SNPs, o trabalho de GONÇALVES et al. (2006) que inclui amostras das populações dos Açores, Madeira, Norte, Centro e Sul do continente Português, refere que o haplogrupo

mais frequente em Portugal é o R1b3 (típico das populações Europeias), que constitui mais de metade das linhagens observadas nas populações referidas; juntamente com os haplogrupos J (típico das populações do Médio Oriente e Norte de África, em particular das populações Judaicas) e E3b (típico do Noroeste e Este Africano e Médio Oriente), é atingida uma percentagem de 78-83% do pool genético de cada população considerada. Os autores referem ainda que metade do haplogrupo E3b2 é partilhado com a população Marroquina, o que sugere a ocorrência de fluxo genético entre as populações de Portugal e Marrocos. Por seu turno ADAMS et al. (2008) reportam a existência de diferenças entre as populações a Norte e a Sul do Rio Mondego, sendo que a Sul observam-se frequências mais elevadas dos haplogrupos E3b2 e J em detrimento do haplogrupo tipicamente Europeu R1b3. Já, BELEZA et al. (2006) referem que ao nível dos distritos sul do Alentejo, foram identificados valores mais elevados de diversidade (número superior de linhagens diferente) do que nas outras províncias, observando-se diferenças significativas relativamente às províncias costeiras do Norte e Centro (Entre Douro e Minho e Beira Litoral). O estudo de CARVALHO et al. (2008) relativamente à população do Centro de Portugal conclui que esta apresenta características em tudo semelhantes às descritas no trabalho de BELEZA et al. (2006) no que diz respeito à diversidade e frequências dos haplogrupos. Em relação à população dos Açores, MONTIEL et al. (2005) afirma que o perfil genético desta população corresponde ao das populações Europeias Ocidentais, verificando-se no entanto diferenças entre o grupo Ocidental, que apresenta frequências dos haplótipos mais representativos (R1, E e J) bastante diferentes das verificadas a nível global do arquipélago. Por seu turno, PACHECO et al. (2005) identificou a ocorrência de 2 haplogrupos (N3 – tipicamente Asiático – e E3a – tipicamente Africano) na população dos Açores que não se encontram representados na população do território continental Português. Também GUSMÃO et al. (2008a) reporta a existência de diferenças em termos de frequências de haplogrupos entre a população Cigana Portuguesa e a restante população do país, sendo que na primeira se observa uma elevada prevalência do haplogrupo J, que contrasta com uma menor prevalência do haplogrupo R1b1c, e ausência do haplogrupo H comparativamente com a restante população. Outras particularidades relativamente à composição genética da população Cigana que são consequência da sua origem na Índia e migração pelo Médio Oriente, Balcãs, Europa e finalmente Península Ibérica são descritas pelos autores, que defendem que o quadro genético desta população é sugestivo da ocorrência de fluxo genético com as diversas populações que partilham o mesmo espaço geográfico das comunidades Ciganas, à semelhança do que GUSMÃO et al. (2010) concluiu com base na análise de marcadores STR autossómicos. Outra população

com base étnica estudada foi a população Judaica de Trás-os-Montes, tendo sido reportadas por NOGUEIRO et al. (2010), algumas diferenças no que diz respeito às frequências de diversos haplogrupos entre esta população e a população geral Portuguesa, nomeadamente a menor frequência do haplogrupo R1b1b2 típico das populações Europeias Ocidentais em contraste com as frequências aumentadas dos haplogrupos J e que refletem a origem no Médio Oriente desta população. Os autores referem também que a população Judaica Portuguesa apresenta um elevado nível de diversidade em conjunto com o facto de que esta população mostra maior semelhança com populações Judaicas Europeias e do Médio Oriente do que com a população Portuguesa geral é indicativo de uma de duas hipóteses: (1) da existência de um número elevado de fundadores ou elevado nível de homogeneidade, o que se explica pelo tamanho estável da comunidade e/ou (2) a ocorrência de introgressões constantes, não pela população circundante, mas por outras populações Judaicas. No entanto, e ainda segundo os autores, o facto de a frequência do haplogrupo R1b1b2 observada na população Judaica Portuguesa ser maior que a observada noutras populações Judaicas é indicativo de um nível significativo de mistura com populações Ibéricas não Judias.

#### *3.4.4. MARCADORES DO MTDNA*

Em relação ao estudo da variabilidade genética observada ao nível do mtDNA LIMA et al. (2004) classificam a população do Norte de Portugal como sendo tipicamente Europeia, dados confirmados posteriormente noutros trabalhos nos quais o haplogrupo H foi referido como o mais frequente (BENTO et al., 2009a, LIMA et al., 2006, LIMA et al., 2008). A diversidade genética do mtDNA é, reportada como sendo menor do que a observada nas populações do Centro da Europa por PEREIRA et al. (2004) ao contrário do que afirma PEREIRA et al. (2000c). Portugal é, no entanto, segundo PEREIRA et al (2005) o País da Europa com mais elevada proporção de ancestralidade Africana ( $3,2 \pm 0,3\%$ ), reportando frequências do haplogrupo U6, típico das populações do Norte de África e haplogrupo L, típico das populações de África a sul do Sahara como sendo as mais elevadas de entre as populações Europeias (2,39% e 3,83%, respetivamente. Os autores fazem referência a factos históricos como a ocupação do território por povos islâmicos do Norte de África e a escravatura moderna para explicar essas diferenças observadas. Foi também

reportada a ocorrência de um haplogrupo característico da população das ilhas Canárias (o U6b1) na população Portuguesa, o que é indicativo de que a contribuição Norte Africana não se restringe ao acontecimento histórico referido; em vez disso, os autores apontam para o contributo de várias populações Norte Africanas para a elevada frequência observada em relação a este haplogrupo. PEREIRA et al., 2000b referem que a distribuição geográfica do haplogrupo U6 se encontra restrita à região Norte de Portugal ao contrário do haplogrupo L que se encontra ubiquamente distribuído pelo território continental. PEREIRA et al. (2010) referem a existência de diferenças genéticas entre três amostras da população Sul (Coruche, Alcácer do Sal e Pias), sendo que em Alcácer do Sal se observou uma elevada frequência de linhagens L em relação à população Portuguesa em geral que estará, segundo os autores, relacionada com o influxo de escravos Africanos nos séculos 15 e 19 e em Pias, observaram-se frequências elevadas dos haplogrupos I, J e T (típicos de populações Mediterrânicas) que dever-se-á, ainda segundo os autores, a contactos com povos Gregos, Fenícios e Cartagineses que estabeleceram importantes redes comerciais no sul da Península.

No que toca à diversidade genética do mtDNA observada nas populações dos arquipélagos Portugueses, BREHM et al. (2003) refere que ambas as populações apresentam valores de diversidade semelhantes mas haplogrupos com diferentes frequências, sendo que na Madeira é notória a influência das populações Africanas a sul do Sahara (com os haplogrupos L1-L3 a atingir os 13% do total de linhagens) que dever-se-á ao fluxo genético proveniente diretamente de África Ocidental; o haplogrupo U6 está presente com moderada frequência na população da Madeira e com frequência reduzida na população dos Açores. O trabalho de SANTOS et al. (2003) relativo à diversidade genética do mtDNA na população dos Açores refere a existência de todos os haplogrupos típicos das populações Europeias e alguns típicos de populações Africanas (U6, M1, L1, L2 e L3) que refletem a origem diversificada dos colonizadores do arquipélago. Ainda segundo os autores, no grupo Central predominam os haplogrupos Europeus enquanto no grupo Ocidental se observa uma distribuição atípica com ausência dos haplogrupos I, J e W, encontrando-se o haplogrupo U sub-representado, sendo que apenas foram identificados os sub-haplogrupos U2 e U6, pouco comuns na Europa mas muito frequentes no Médio Oriente e Norte de África, respetivamente. O haplogrupo mais comum no grupo Ocidental é o V, ao contrário do que é comum nas populações Europeias (H); esta diferenciação dever-se-á, de acordo com o mesmo trabalho, ao processo de deriva genética induzido pelo reduzido tamanho populacional efetivo e, eventualmente, recente redução

do tamanho da população. Em relação aos grupos Oriental e Central, observou-se uma contribuição diferencial das linhagens Africanas e do Médio Oriente, respetivamente.

Em suma, o perfil da população Portuguesa é classificado como homogéneo e tipicamente Europeu relativamente à generalidade dos marcadores autossómicos e do cromossoma X (BRANCO et al., 2006b, CARVALHO et al., 2011, GUSMÃO et al., 2009, PEREIRA et al., 2008, PEREIRA et al., 2007, PINHEIRO et al., 2005). Relativamente aos marcadores STR do cromossoma Y, observa-se uma elevada diversidade de haplótipos, sendo de destacar a ocorrência de um haplótipo típico do Norte da Europa na população Madeirense que não foi observado na Península Ibérica (FERNANDES et al., 2002b), a diversidade aumentada reportada na população do Alentejo (BELEZA et al., 2006) e a diversidade diminuída reportada na população Cigana Portuguesa (GUSMÃO et al., 2008a) comparativamente ao observado a nível nacional. No que toca à caracterização genética da população com recurso a marcadores SNP do cromossoma Y, observa-se um predomínio do haplogrupo tipicamente Europeu R1b3 (cromossoma Y) (GONÇALVES et al., 2006), tendo sido Foram também reportadas diferenças regionais em termos de diversidade, sendo que se observaram níveis de diversidade mais elevados no Alentejo do que no restante território continental (BELEZA et al., 2006), e a ocorrência de dois haplogrupos (N3 e E3a) na população dos Açores que não se encontram descritos nas populações do território continental Português (PACHECO et al., 2005). Foram também observadas diferenças em termos de frequências e diversos haplogrupos entre a população geral Portuguesa e as populações Cigana (GUSMÃO et al., 2008a) e Judaica (NOGUEIRO et al., 2010). Em relação à diversidade observada ao nível do mtDNA, apesar de se observar uma elevada frequência dos haplogrupos Africanos U6 e L relativamente à generalidade dos países Europeus (PEREIRA et al., 2005), o haplogrupo H (típico de populações Europeias) é o mais frequente (BENTO et al., 2009a, LIMA et al., 2006, LIMA et al., 2008), com exceção do grupo Ocidental do arquipélago dos Açores, população na qual o haplogrupo mais frequente é o V (SANTOS et al., 2003). Na população da Madeira é observado um aumento das frequências de haplogrupos Africanos L1-L3 e U6 (BREHM et al., 2003), sendo que o último se encontra, ao nível do território continental, restrito à população Norte (PEREIRA et al., 2000b).





# CAPÍTULO 4

## 4. QUESTÕES BIOÉTICAS E LEGAIS

A previsão de traços físicos com base na análise de marcadores genéticos quer pelo método direto quer indireto, tem o potencial de fornecer pistas importantes no âmbito da investigação forense. No entanto, a aplicação destes métodos de análise levantam questões de ordem ética e legal cuja descrição se segue.

### 4.1. QUESTÕES BIOÉTICAS

Apesar do objetivo de utilização dos métodos de previsão, tanto do método direto como o indireto, ser a previsão da aparência de dadores de amostras em casos nos quais esses mesmos dadores são pessoas não identificadas, uma das principais questões legais relacionadas com esta matéria é a questão da violação da privacidade. KAYSER et al. (2009) argumenta, neste sentido, que “não existe violação do direito de privacidade na medida em que a informação obtida não se encontra ligada a uma pessoa específica”, referindo ainda que “as características exteriormente visíveis de uma pessoa são visíveis para todos, pelo que a informação que a previsão genética [dessas características] revela não pode ser considerada como informação privada”. Para além disso, os kits de marcadores autossómicos de identificação forense incluem um marcador que permite prever o sexo do dador das amostras (localizado no gene da Amelogenina), sendo essa informação incluída nas bases de dados de identificação.

Por outro lado, os autores afirmam que não se coloca o problema da proteção de dados uma vez que estes são apenas interessantes no âmbito de cada caso em particular, não existindo, portanto, a necessidade de armazenar os dados de previsão em bases de dados. Os autores enfatizam, também, o facto de que a informação obtida através do recurso a marcadores genéticos de previsão permitir apenas prever a probabilidade da manifestação de uma determinada característica e não ter a certeza de que essa característica se manifesta de facto no indivíduo. Assim, a análise estatística dos dados obtidos permite apenas reduzir o número de potenciais suspeitos e não identificar o dador do material biológico em concreto, cabendo esse papel aos marcadores de identificação. Neste sentido, a informação obtida através do estudo de marcadores preditivos de DNA assemelha-se à fornecida por testemunhas oculares (FRUDAKIS, 2008, KAYSER et al., 2009).

Em relação às identificações realizadas por testemunhas oculares há alguns fatores a ter em consideração, em particular, os procedimentos seguidos, como por exemplo, os alinhamentos (Figura 56), que, segundo SPINNEY (2008), podem ter alguma interferência na própria identificação. A autora refere que a forma como o procedimento é levado a cabo tem influência na testemunha que pode ser levada a identificar um suspeito apenas porque percebe, através da “linguagem corporal, tom de voz ou outras dicas inconscientes” do inspetor, qual o suspeito entre os indivíduos presentes no alinhamento, ao contrário do que se espera, que é a identificação com base na recordação dos acontecimentos.



**Figura 56 – Alinhamento de suspeitos para identificação. Adaptado de SPINNEY (2008).**

A questão dos falsos testemunhos está também referida na publicação do INNOCENCE PROJECT (2010), grupo de advocacia fundado em 1992 na Benjamin N. Cardozo School of Law da Universidade de Yeshiva em Nova Iorque, que auxilia prisioneiros cuja inocência pode ser comprovada através de análises de DNA. A referida publicação intitulada *“250 Exonerated, Too Many Wrongfully Convicted”* (250 Exonerados, Demasiados Condenados Injustamente) relata os primeiros 250 casos de exonerações de indivíduos condenados cuja inocência foi provada através de testes de identificação de DNA. É importante salientar algumas das conclusões deste trabalho, nomeadamente o facto de 76% desses indivíduos terem sido condenados com base, pelo menos em parte, em identificações erradas de testemunhas oculares, sendo que 53% dessas identificações erradas foram relativas a indivíduos de outras raças (na maioria dos casos “uma pessoa brancas que identifica uma pessoa negra”). Outros dados interessantes deste trabalho são referentes às consequências da condenação de pessoas inocentes. Por um lado, a vida dessa pessoa é radicalmente e injustamente alterada (tempo médio de cativeiro foi de 13 anos, tendo atingido os 35 anos num dos casos), correndo até o risco de ser condenada à morte (17 dos 250 casos são exemplo disso mesmo). Por outro lado, os verdadeiros perpetradores dos crimes permanecem à solta dando continuidade à prática criminosa, sendo que, ainda de acordo com o mesmo trabalho, “caso o verdadeiro perpetrador tivesse sido originalmente condenado, em vez de uma pessoa inocente, pelo menos 72 crimes violentos poderiam ter sido prevenidos”.

Por seu turno, KAYSER et al. (2009) defendem que a previsão da aparência do suspeito através do estudo de marcadores genéticos, por um lado assenta em dados objetivos (probabilidades) sendo, por isso, neutra e, por outro, dependendo da característica em análise, permite obter informações tão ou mais fiáveis do que as fornecidas pelas testemunhas. Também FRUDAKIS (2008) faz referência às vantagens do uso do sistema de reconhecimento facial reverso (anteriormente descrito) realizado a partir do estudo de marcadores genéticos de previsão comparativamente à realização de esboços do rosto de suspeitos com base no relato de testemunhas oculares que “frequentemente estão equivocadas ou são desonestas ou simplesmente imprecisas”. O autor reconhece, no entanto, a importância da “capacidade inata que todos os seres humanos têm de reconhecer faces e perceber semelhanças e pequenas diferenças entre indivíduos”. Trata-se de uma função biológica de elevada importância na medida em que, em todas as culturas “constitui um requisito básico para a interação social eficiente” (CALDARA et al., 2010). Esta capacidade é, no entanto, influenciada por fatores étnicos,

sendo, por exemplo, “mais difícil para um Europeu discriminar entre rostos Asiáticos do que para um Asiático” (FRUDAKIS, 2008). Este fenómeno psicológico de variação nos níveis de sensibilidade para determinados componentes de ancestralidade são influenciados pela experiência cultural e designa-se “propensão da própria raça” ou “efeito de outra raça” LINDSAY et al. (1991) citado por FRUDAKIS (2008). LEBRECHT et al. (2009) refere, no entanto, que este fenómeno psicológico é multifacetado e, por incluir capacidades percetivas moldáveis, pode ser alterado com algum treino.

No âmbito da previsão direta das características fenotípicas são analisadas as bases genéticas subjacentes a diversos componentes da aparência dos seres humanos. Teme-se que este conhecimento possa ser usado por casais que pretendem procriar para selecionar as características da descendência com base na aparência (GRAHAM, 2008). Esta seleção não natural da descendência põe em causa a diversidade genética da espécie que é garantida pelos mecanismos aleatórios inerentes ao processo de procriação. Por outras palavras, em vez de os seres humanos gerados serem o produto da fecundação que junta gâmetas aleatórios, cada um com um conjunto aleatório de genes dos progenitores, passar-se-ia a gerar embriões *in vitro* que seriam testados para que a sua aparência fosse prevista com base nos seus dados genéticos e seriam selecionados “os mais bonitos”, de acordo com as preferências dos pais.

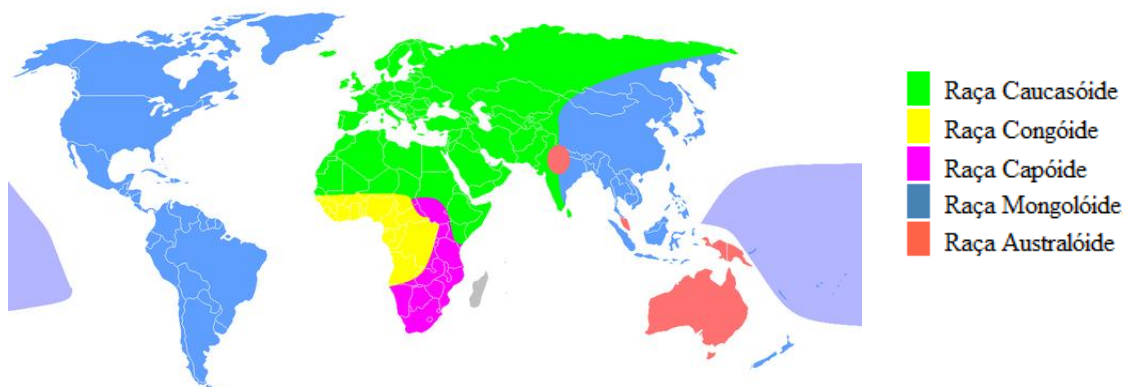
A ideia de que determinadas características físicas definem indivíduos superiores era defendida por Hitler e serviu de argumento ao extermínio de indivíduos considerados “inferiores”.

Outra questão que se coloca no âmbito da utilização indevida da informação obtida a partir de marcadores de ancestralidade é colocada por um grupo multidisciplinar da Universidade de Stanford: “sob que condições poderão os dados genéticos transformar os entendimentos sociais de categorias raciais e étnicas, eventualmente exacerbando ideologias racistas?” (LEE et al., 2008). OBASOGIE (2009) afirma que nas últimas décadas, a mudança de pensamento sobre o conceito de “raça” como categoria em função de diferenças para reflexo de escolhas sociais e culturais constituiu a base do progresso das relações raciais. Neste sentido, o autor defende que as “noções biológicas de diferença racial” poderão reverter o progresso observado “assim que estas voltem a estar incluídas nas conversas científicas e públicas”.

A classificação da espécie humana em raças ou subespécies é um dos assuntos que levanta maior controvérsia no seio da comunidade científica. Na Tabela 14 encontram-se descritos sistemas de classificação racial citados por TISHKOFF et al. (2004). Os autores referidos propõem diferentes formas de classificar os seres humanos. A 1ª classificação de Linnaeus classifica em subespécies e, posteriormente, o autor adicionou subdivisões geográficas, (2ª classificação) que têm por base dados biológicos e culturais. Blumenbach, por seu turno, propôs uma classificação racial que inclui os grupos descritos por Linnaeus mais o grupo “Malay” que corresponde aos indivíduos que habitam a Oceânia e, finalmente, o antropólogo Carleton Coon propôs um sistema de classificação racial com base em características fenotípicas físicas, que, apesar de controverso, é, ainda de acordo com TISHKOFF et al. (2004) nos dias de hoje muito utilizado (Tabela 14 e Figura 57). JACKSON JR (2001), por seu turno, defende que a divisão em cinco raças da espécie humana que Coon propôs constituiu uma “violação aos seus próprios padrões de objetividade científica”, servindo interesses do movimento segregacionista nos EUA, que defendia que os cidadãos negros tinham um estatuto inferior aos cidadãos brancos.

**Tabela 14 – Principais sistemas de classificação racial com base em TISHKOFF et al. (2004).**

Autor	Classificação	
Linnaeus (1758)	1ª classificação	<ul style="list-style-type: none"> <li>• <i>Homo sapiens afer</i></li> <li>• <i>Homo sapiens europaeus</i></li> </ul>
	2ª classificação	<ul style="list-style-type: none"> <li>• Europeus brancos</li> <li>• Americanos vermelhos</li> <li>• Asiáticos amarelos</li> <li>• Africanos negros</li> </ul>
Blumenbach (1775)	Adicionou o grupo “Malay” (indivíduos da Oceânia) à 2ª classificação de Linnaeus	
Coon (1962)	<ul style="list-style-type: none"> <li>• Raça Caucasóide (Europa, Oeste da Ásia e Norte de África)</li> <li>• Raça Mongolóide (Este da Ásia, Ártico e Américas)</li> <li>• Raça Australóide (Oceânia e Melanésia)</li> <li>• Raça Negróide (África a Sul do Sahara)</li> <li>• Raça Capóide (Sul da África)</li> </ul>	



**Figura 57 – Mapa no qual se encontra representada a distribuição geográfica das raças humanas definidas por Carleton Coon. Figura adaptada de [http://en.wikipedia.org/wiki/File:Carleton\\_Coon\\_races\\_after\\_Pleistocene.PNG](http://en.wikipedia.org/wiki/File:Carleton_Coon_races_after_Pleistocene.PNG), consultado a 07-09-2011.**

Charles Darwin, o pai do Evolucionismo, por seu turno, apesar de ter observado diferenças entre seres humanos, nomeadamente ao nível do tom da pele, defendia que esse parâmetro não era uma boa base de distinção entre espécies ou raças humanas por apresentar uma variação gradual. A sua posição foi certamente influenciada pelo período histórico em que viveu, período no qual a escravatura (prática social em que um ser humano é privado da liberdade (escravo), sendo submetido a um poder absoluto por via da força por outro ser humano que assume sobre o primeiro direitos de propriedade (Infopédia - Enciclopédia e Dicionários Porto Editora)), era prática comum, sendo que Darwin não concordava com esses mesmos ideais (JABLONSKI et al., 2010).

De facto, os dados genéticos apontam para o facto de que as populações humanas se agrupam, de modo geral, em extensas regiões geográficas – África a sul do Sahara; Europa, norte de África e oeste da Ásia; este da Ásia; Oceânia e, finalmente, Américas (TISHKOFF et al., 2004). No entanto, segundo LEE et al. (2008), “as categorias raciais e étnicas são criadas e mantidas dentro de determinados contextos sociais e políticos” não correspondendo muitas vezes às categorias de ancestralidade genética. Observa-se, portanto, uma correlação imperfeita entre ancestralidade biogeográfica e grupos étnicos ou raciais devido ao facto de a variação genética ser tendencialmente contínua com alguma sobreposição entre as populações (JORDE et al., 2004). Assim, a ancestralidade da maioria dos indivíduos é composta pela mistura de diversas proporções de dois ou mais grupos ancestrais FRUDAKIS (2008), pelo que não é consensual a utilização de designações raciais na construção de perfis de ancestralidade biogeográfica (RACE

ETHNICITY AND GENETICS WORKING GROUP, 2005)<sup>3</sup>. Um exemplo dado por LEE et al. (2008) são os Hispânicos, grupo muito heterogêneo de indivíduos com diferentes ancestralidades e histórias sociais. O mesmo grupo interdisciplinar enfatiza o facto de que a sobrevalorização da contribuição dos fatores genéticos para a determinação de características complexas como o QI, a tendência para a violência e a performance atlética promove um essencialismo genético ingênuo – a noção de que os genes determinam o comportamento e eventualmente ideologias racistas.

Outro aspeto importante apontado por LEE et al. (2008) consiste na inexistência de base científica para a hierarquização de categorias raciais ou étnicas.

Por outro lado, a interpretação dos resultados obtidos através da análise do perfil de ancestralidade pode ser influenciada por atitudes sociocognitivas (preconceito racial) relativamente a outros grupos étnicos ou populações, estando relacionada com o “Efeito da Outra Raça” que implica uma maior dificuldade no reconhecimento facial de indivíduos de outras raças. Esse efeito pode, no entanto, ser minorado através de treino, que indica que o preconceito racial inclui uma componente perceptual que pode ser modificada (LEBRECHT et al., 2009).



**Figura 58 – Dois exemplos do passado de situações extremas que constituíram consequências do domínio de ideais racistas: a subjugação de seres humanos – escravatura (à esquerda) – e o extermínio de um povo – Holocausto (à direita). Figura da esquerda adaptada de: <http://slaveryresistancewestciv.blogspot.com/2011/05/slave-brutality.html> e figura da direita adaptada de: <http://my.opera.com/BlueMike/blog/show.dml/7005091>, consultados a 07-09-2011.**

---

<sup>3</sup> O *Race, Ethnicity and Genetics Workgroup* corresponde a um grupo de trabalho do *National Human Genome Research Institute* (Estados Unidos da América) formado com o intuito de abordar as questões relacionadas com a raça e genética). *Genome Research Institute* (Estados Unidos da América) formado com o intuito de abordar as questões relacionadas com a raça e genética).

O RACE ETHNICITY AND GENETICS WORKING GROUP (2005) acrescenta, ainda, que se por um lado o estudo das diferenças interpopulacionais e da sua relação com traços físicos e até risco de doenças poderá reforçar estereótipos existentes, por outro constitui uma fonte de informação relativa às origens complexas dessas mesmas características e à proximidade biológica entre os grupos humanos, o que, em sentido contrário, retira fundamento a esses mesmo estereótipos.

O Holocausto, a escravatura, os conflitos com base étnica ou racial têm lugar desde muito antes de se estudarem os marcadores genéticos relacionados com essas diferenças e, de acordo com GUIMARÃES (2005), “foi a adoção de uma visão equivocada da biologia humana, expressa pelo conceito de ‘raça’, que estabeleceu uma justificativa para a subordinação permanente de outros indivíduos e povos, temporariamente sujeitos pelas armas, pela conquista, pela destituição material e cultural, ou seja, pela pobreza”. Por seu turno, FRUDAKIS (2008) coloca a questão da seguinte forma: “O que teria feito Hitler com um teste de ancestralidade?”, respondendo o mesmo autor à questão referindo que todas as novas tecnologias apresentam potencial uso inapropriado (os aviões podem ser usados em atentados e os medicamentos, facas ou até lápis em homicídios), pelo que esse risco de má utilização está, obviamente presente; o uso inapropriado não é, no entanto, consequência da disponibilidade de ferramentas, mas sim do comportamento humano, das decisões que as pessoas tomam. Neste sentido, LEE et al. (2008) recomenda a “inclusão de informação científica social e histórica acerca de usos da ciência no passado que promoveram o racismo bem como o potencial impacto em políticas futuras” no ensino da genética como forma de prevenir a repetição de acontecimentos como o Holocausto, massacre, sobretudo de Judeus, que teve lugar na Segunda Guerra Mundial nos campos de concentração nazis, massacre esse justificado por ideais de superioridade/inferioridade entre raças/etnias humanas?

#### 4.4. LEGISLAÇÃO INTERNACIONAL

A Holanda é o único país no mundo no qual é, em termos legais, explicitamente permitida e regulamentada a prática da fenotipagem de DNA em contexto forense (KAYSER et al., 2011, KAYSER et al., 2009, KOOPS et al., 2008). Noutros países (Bélgica e Alemanha) e alguns Estados americanos (Indiana, Rhode Island e Wyoming) é



expressamente proibida a aplicação de testes de DNA com o objetivo de obter informações relativas à aparência dos indivíduos (com exceção do sexo) (KOOPS et al., 2008).

Na maioria dos países a legislação é omissa no que toca à previsão de características fenotípicas a partir de amostras de DNA o que torna possível a sua utilização por não haver uma proibição explícita. Este facto deve-se à ausência de conhecimento relativo a esta área, às limitações técnicas e também às questões éticas subjacentes que tornam imperativa a discussão em praça pública relativamente à sua utilização ou não bem como as regras de utilização (KAYSER et al., 2009, KOOPS et al., 2008). Assim, segundo KOOPS et al. (2008), “dependendo do sistema legal e cultura, [a fenotipagem de DNA] pode ser implicitamente permitida como parte das investigações forenses” dando como exemplos o Reino Unido e Japão.

#### 4.5. EM PORTUGAL

Como foi anteriormente referido, em Portugal, não são realizadas análises de DNA no contexto de previsão de características fenotípicas, tendo o Dr. Bruno Antunes<sup>4</sup> referido que a não utilização destes métodos de análise se deve parcialmente a limitações legais. De facto, não existe legislação específica que proíba ou permita e regule a utilização de marcadores de previsão de fenótipo e ancestralidade com aplicação forense a nível nacional de acordo com a Prof. Doutora Helena Moniz<sup>5</sup> e a Dr.<sup>a</sup> Ana Luísa Zêzere<sup>6</sup>, pelo que se segue a referência a disposições legais relacionadas com a análise de DNA em contexto forense bem como outras normas e pareceres relacionados com a potencial aplicação de marcadores de previsão de fenótipo e ancestralidade e também com o a nível nacional.

A Lei nº 5/2008 de 12 de Fevereiro que aprova a criação de uma base de dados de perfis de ADN para fins de identificação civil e criminal estipula no artigo 12.º (Âmbito da análise): “1 — A análise da amostra restringe-se apenas àqueles marcadores de ADN que sejam absolutamente necessários à identificação do seu titular para os fins da presente lei. 2 — Os marcadores de ADN a integrar no ficheiro de perfis de ADN são fixados, após

---

<sup>4</sup> Dr. Bruno Antunes, Perito de Recolha de Vestígios no Local de Crime do Laboratório de Polícia Científica da Polícia Judiciária – comunicação pessoal.

<sup>5</sup> Prof. Doutora Helena Moniz, Professora Auxiliar da Faculdade de Direito da Universidade de Coimbra – comunicação pessoal.

<sup>6</sup> Dra. Ana Luísa Zêzere, Procuradora do Ministério Público – comunicação pessoal.

parecer da Comissão Nacional de Proteção de Dados (CNPd), por portaria (...)”. Na alínea e) do artigo 2.º encontra-se a definição de marcador de ADN: “região específica do genoma que tipicamente contém informações diferentes em indivíduos diferentes, que segundo os conhecimentos científicos existentes não permite a obtenção de informação de saúde ou de características hereditárias específicas, abreviadamente ADN não codificante”.

Na Portaria nº 270/2009 de 17 de Março encontram-se estabelecidos os marcadores que integram a base de dados nacional. Relativamente ao conteúdo informativo desses marcadores, o nº1 do artigo 11.º do Regulamento de funcionamento da base de dados de perfis de ADN (publicado no âmbito da Deliberação n.º 3191/2008 do Instituto Nacional de Medicina Legal em Diário da República, 2.ª série – N.º 234 – 3 de Dezembro) prevê que “no caso de algum dos marcadores de ADN revelar informação relativa à saúde ou a características hereditárias específicas, esse marcador é excluído dos perfis de ADN incluídos na base de dados e deixa de ser estudado nas amostras a analisar posteriormente”. O nº 2 do mesmo artigo refere ainda a exclusão do género relativamente à norma anterior.

A interdição descrita nas normas anteriormente referidas relativamente ao uso de marcadores codificantes servirá de fronteira entre o que se pode ou não analisar sob o ponto de vista forense e incluído na base de dados nacional. No entanto, o genoma humano encontra-se organizado de forma muito particular (como foi já abordado), sendo que o DNA é herdado em “blocos” (haplótipos) praticamente intactos constituídos por material codificante e não-codificante. Isto significa que a análise de um marcador não-codificante que se encontre fisicamente próximo (no mesmo haplótipo) do marcador codificante revela a mesma informação uma vez que estes se encontram em desequilíbrio de ligação (KAYSER et al., 2009).

Pode ainda acrescentar-se, em relação à base de dados nacional de DNA, que as regras de inclusão de perfis (descritas na Lei nº 5/2008 de 12 de Fevereiro) são bastante restritivas, sendo que essa inclusão (artigo 15º), quando não é realizada a pedido do indivíduo (voluntário, arguido ou familiar de pessoa desaparecida) pode ser realizada através de despacho de juiz no âmbito de um processo crime ou “após trânsito em julgado, (...) [tendo sido o réu] condenado por crime doloso com pena concreta de prisão igual ou superior a 3 anos” (nº2 do artigo 8º). No último caso, o perfil permanece na base de dados até à “data em que se proceda ao cancelamento definitivo das respectivas decisões no

registo criminal” (alínea *f* do n.º 1 d artigo 26º). Estas restrições traduzem-se, na prática, num reduzido número de perfis presente na base de dados na qual constam 286 perfis<sup>7</sup>.

A legislação citada até agora referida diz apenas respeito aos dados a incluir na base de dados nacional. Como o objetivo da análise de marcadores de previsão não é a criação de bases de dados nem identificar inequivocamente o dador de uma amostra problema, mas sim prever características da aparência do dador da amostra problema com o objetivo de restringir o número de suspeitos a considerar no âmbito de uma investigação, torna-se pertinente a referência a outros diplomas legais potencialmente relevantes.

A Constituição da República Portuguesa, por exemplo, reconhece a todos os cidadãos nacionais o “direito à identidade pessoal, (...) à imagem, (...) e à proteção legal contra quaisquer formas de discriminação” (nº 1 do Artigo 26.º Capítulo I).

Segundo o ponto 1 do Artigo 7º da Secção I do Capítulo II da Lei nº 67/98 de 26 de Outubro – Lei da Proteção de Dados Pessoais – “É proibido o tratamento de dados pessoais referentes a (...) origem racial ou étnica (...)”. No entanto o ponto 2 do mesmo artigo refere que “Mediante disposição legal ou autorização da CNPD, pode ser permitido o tratamento dos dados referidos no número anterior quando por motivos de interesse público importante esse tratamento for indispensável ao exercício das atribuições legais ou estatutárias do seu responsável (...)”. De facto, dados relativos à origem étnica do indivíduo e progenitores são incluídos na base de dados nacional de DNA. De acordo com o Regulamento de funcionamento da base de dados de perfis de ADN, esses mesmos dados constam nos Autos de colheita de amostras para identificação civil (Anexo II-B) e nos Autos de colheita de amostras e de identificação em voluntários (Anexo II-A), em condenados (Anexo II-C) e em arguídos (Anexo II-D).

É também importante referir alguns pareceres do Concelho Nacional de Ética para as Ciências da Vida, cuja primeira competência é a de “analisar sistematicamente os problemas morais suscitados pelos progressos científicos nos domínios da biologia, da medicina ou da saúde em geral”, de acordo com o nº1º do Artigo 2º da Lei nº 14/90 de 9 de Junho. Assim, o “Relatório e Parecer Sobre Implicações Éticas da Genómica” (CONSELHO NACIONAL DE ÉTICA PARA AS CIÊNCIAS DA VIDA, 2001) constitui uma reflexão sob o ponto de vista ético acerca das potenciais aplicações do conhecimento do genoma humano. São abordadas questões como a influência dos fatores genéticos na

---

<sup>7</sup> Dados fornecidos pelo Prof. Doutor Francisco Corte Real, Director da Delegação Centro do Instituto Nacional de Medicina Legal a 29-11-2011 – comunicação pessoal.

determinação do “espírito humano” (comportamento), escolha das características da descendência, regras de aplicação de testes genéticos preditivos (no âmbito da previsão de predisposição a doenças genéticas), investigação científica em seres humanos e embriões humanos entre outras. Os pareceres que constam na parte final do documento apontam para a necessidade de incentivar as investigações sobre o genoma humano por constituírem uma importante fonte de conhecimento da origem evolutiva do ser humano e permitirem também compreender a influência relativa e não absoluta dos genes nas características humanas ou, por outras palavras a “margem de liberdade humana face à ação dos genes”. Também a “luta contra o racismo” é referida como uma das “consequências dessas investigações para o conhecimento e a imagem que o ser humano vai formando de si próprio”. É também enfatizada a importância da “proteção da confidencialidade dos dados genéticos” e do consentimento informado em relação aos testes genéticos preditivos. É considerada inaceitável, sob o ponto de vista ético, a seleção das características físicas da descendência. Existem, no entanto, um enquadramento legal que prevê realização do diagnóstico genético pré-implantação em casos particulares (lei nº 32 de 2006 de 26 de Julho – Procriação medicamente assistida), nomeadamente casos de “pessoas provenientes de famílias com alterações que causam morte precoce ou doença grave quando exista o risco elevado de transmissão à descendência” n.º 1 do artigo 29º do Capítulo V).

No parecer nº 52 – Parecer sobre o regime jurídico da base de dados de perfis de ADN, emitido a 12 de Junho de 2007, consta na última alínea: “qualquer alteração do âmbito de uma base de perfis de ADN, construída inicialmente com fins de investigação criminal, para outras finalidades, novas utilizações e critérios de inclusão, deverá ser sujeita a amplo debate público” (CONSELHO NACIONAL DE ÉTICA PARA AS CIÊNCIAS DA VIDA, 2007).

## DISCUSSÃO, CONCLUSÕES E PERSPETIVAS FUTURAS

Os marcadores genéticos apresentam diferentes características consoante a sua localização no genoma e também o tipo de polimorfismo analisado. Os marcadores de identificação permitem proceder à identificação do dador de uma amostra com base na comparação de perfis. Atualmente os marcadores mais usados na identificação são os STR mas existe já um painel de SNPs para o mesmo fim. Na mesma linha, os SNPs, por serem menos polimórficos e mais estáveis que os STRs mantendo-se inalterados durante várias gerações (por exemplo ao nível do mtDNA e cromossoma Y), representam o tipo de polimorfismos mais utilizados na inferência de informações relativas à ancestralidade biogeográfica. Também o facto de se localizarem frequentemente em regiões codificantes faz com que muitas vezes estejam relacionados com alterações fenóticas, pelo que são muito usados em estudos de associação. Assim, no âmbito da previsão de fenótipos pelo método direto e indireto os polimorfismos de eleição são os SNPs, o que não exclui a utilização de outros marcadores. O critério para que um determinado marcador seja considerado informativo relativamente à ancestralidade de um indivíduo (por outras palavras, para que seja um AIM) é que exiba variação em termos de frequências alélicas em diferentes populações, refletindo o efeito das forças evolutivas na moldagem da composição genética das mesmas.

A localização dos marcadores é importante na fase de interpretação de resultados, uma vez que marcadores autossómicos e marcadores de linhagem (NRY e mtDNA) produzem diferentes tipos de informação. A análise de marcadores de linhagem permite apenas aceder a uma pequena porção da informação genética do indivíduo que se mantém praticamente inalterada ao longo de várias gerações. A informação fornecida pelos marcadores autossómicos, essa sim, reflete todo o espectro de ancestralidade biogeográfica do indivíduo que é o produto da contribuição de todos os seus antepassados.

Os progressos técnicos e científicos permitem atualmente prever, de forma direta ou indireta, com maior ou menor precisão algumas características físicas e essas informações já se revelaram bastante úteis na restrição do número de indivíduos a considerar no

âmbito da investigação de crimes ou até identificação de cadáveres em avançado estado de decomposição. As limitações técnicas são ainda muitas, particularmente no que diz respeito à precisão da previsão direta relativamente à maioria das características exceto dos traços relacionados com a pigmentação, estando já identificados alguns conjuntos de um número reduzido de marcadores que permitem prever com bastante precisão (acima de 80% para a cor do cabelo e de 90% para a cor dos olhos) estas características. Em relação aos três painéis de previsão analisados constatou-se a partilha de marcadores para prever diferentes características. Esses marcadores partilhados são reportados noutros trabalhos como loci que apresentam variação interpopulacional. Também a revisão dos trabalhos relativos aos genes e respetivos polimorfismos relacionados com os traços de pigmentação, permitiu identificar, em relação aos sete polimorfismos mais citados nesse conjunto de estudos, frequências alélicas diferentes em diferentes populações. Em relação aos genes mais citados, em todos eles se identificou pelo menos um polimorfismo nas mesmas condições. Estes factos são indicativos, por um lado, da forte relação entre a variação fenotípica observada relativamente aos traços de pigmentação e a origem biogeográfica e, por outro, de que a associação que é estabelecida entre o genótipo e o fenótipo poderá não ser causal, ou seja, a variação genotípica está relacionada com a ancestralidade biogeográfica e não diretamente com a característica em particular. Por outro lado é de realçar que a maioria dos trabalhos tem por base a população Europeia, na qual se observa grande variabilidade intrapopulacional ao nível da cor dos olhos e do cabelo comparativamente a outras populações. Em relação à altura e traços faciais, os resultados dos estudos de associação são ainda bastante limitados (relativamente à percentagem da variação fenotípica que se consegue explicar pela variação nos marcadores analisados), indicando a existência de um grande número de loci envolvidos na determinação dessas mesmas características. Assim, é previsível que o número de marcadores necessários para previsão directa destes traços seja elevado, dificultando a aplicação do método direto na prática quando todos os loci forem conhecidos.

A variação observada em traços como a morfologia capilar, índice de massa corporal e traços faciais exhibe uma correlação com a ancestralidade biogeográfica que indica que, a curto prazo, a previsão pelo método indireto é, também em relação a estes traços, a melhor opção. A informação obtida através do método indireto permite prever a aparência do dador da amostra em análise através da relação existente entre a ancestralidade biogeográfica e a variação genotípica, no entanto, a bibliografia sugere que é preferível a criação de bases de dados para que as correlações estabelecidas entre

ancestralidade biogeográfica e as características físicas sejam realizadas de forma objetiva e seja possível quantificar essa correlação. Essas bases de dados não têm o propósito de armazenar informação relativa a casos forenses, mas sim informação relativa à população.

O método indireto poderá também servir de base a estudos de associação para a identificação das bases genéticas que se relacionam diretamente com os traços referidos, sendo previsível a substituição deste pelo método direto, uma vez que o último permite a realização de previsões mais precisas, visto que se baseia na análise dos polimorfismos diretamente relacionados com os traços e não implica a utilização de bases de dados.

Para além das características relacionadas com a pigmentação, também o sexo e a idade são características que é possível prever de forma direta, embora, no primeiro caso tenham sido reportadas taxas de erro que são indicadoras da necessidade de utilizar outros marcadores, no caso da altura, por se tratar de uma característica determinada pela ação de um grande número de loci, não se vislumbra num futuro próximo a possibilidade de se prever esta característica a partir do DNA. Em relação à idade, o método com melhores resultados (quantificação de moléculas de DNA circular nas células T) carece ainda de estudos mais aprofundados sobre eventuais fatores que não tenham ainda sido considerados para que o intervalo de previsão (atualmente de cerca de 20 anos) diminua.

Outro fator a ter em conta relativamente à previsão de características que constituem a aparência de um dador de uma amostra de DNA é o facto de que essas características mudam durante a vida do indivíduo (o tom de pele e a cor do cabelo, por exemplo), sofrem alterações decorrentes da ação de fatores ambientais (a exposição solar torna a pele mais escura e oxida o cabelo, tornando-o mais claro, por exemplo) e podem até ser modificados voluntariamente (coloração de cabelo, lentes de contacto coloridas, cirurgia de mudança da cor dos olhos e de mudança de sexo, por exemplo). Este fator é indicativo do valor relativo que caracteriza a previsão de características físicas com base no DNA e da forma cuidadosa como os dados de previsão devem ser encarados.

Embora existam vários grupos de trabalho por todo o mundo a desenvolver estudos nesta área, é possível destacar dois nomes (que representam dois grupos de trabalho) que têm contribuído de forma mais significativa para o crescente conhecimento que se tem vindo a verificar na área da previsão de características físicas no âmbito da análise forense: Tony Frudakis e Manfred Kayser. Tony Frudakis foi um dos fundadores da DNAPrint Genomics, empresa Norte Americana que entretanto cessou a atividade em 2008

(GENOMEWEB DAILY NEWS, 2009, THE GENETIC GENEALOGIST, 2009). A referida empresa continua, no entanto, a licenciar o DNAWitness a laboratórios privados e um dos co-fundadores, Mark Shriver, continua a desenvolver investigação na área da previsão de características fenotípicas na Pensilvânia (FOX, 2010). Por seu turno, Manfred Kayser lidera o Departamento de Biologia Molecular Forense da Universidade Erasmus (Roterdão – Holanda) dedicando-se ao estudo da ancestralidade biogeográfica e características fenotípicas, entre outras áreas de aplicação forense.

No que toca aos métodos de fenotipagem, alguns autores (incluindo o grupo de Frudakis) privilegiam os métodos de avaliação quantitativa ao contrário de outros (como por exemplo o grupo de Kayser) que recorre normalmente a métodos de avaliação por observação (auto e heterodescrição). Estas diferentes abordagens poderão ser influentes nos resultados obtidos nos testes de associação.

A revisão de trabalhos relativos à caracterização da variação genética da população portuguesa permitiu concluir que o perfil genético da população nacional é bastante homogéneo relativamente aos marcadores autossómicos e do cromossoma X. Uma vez que se trata de estudos relativos a marcadores de identificação, essa caracterização é esperada, pois são marcadores altamente polimórficos em todas as populações. Como contraste surge a caracterização genética dos marcadores de linhagem, nos quais se observa alguma diferenciação entre populações geograficamente definidas, salientando-se as regiões dos Açores, Madeira e Alentejo a apresentar algumas particularidades em termos de diversidade e até frequências alélicas comparativamente às restantes regiões. Foram também analisados os resultados de estudos relativos à caracterização genética de grupos étnicos (Ciganos e Judeus) relativamente aos quais foram observadas diferenças genéticas ao nível dos marcadores de linhagem comparativamente à população geral do país que decorre do relativo isolamento destes grupos. Não se encontram, porém, descritos estudos relativos à aplicação de painéis de AIMs autossómicos na população portuguesa.

No que diz respeito às questões de ordem ética e legal, por um lado, e como foi já referido, a previsão de características físicas a partir da análise de DNA constitui uma ferramenta com um enorme potencial de utilidade na investigação criminal, em casos nos quais se dispõe de amostras biológicas mas não se conseguiu identificar o dador das mesmas através dos tradicionais testes de identificação; por outro, as vozes críticas relativamente ao desenvolvimento e aplicação desta tecnologia revelam preocupação relativamente às consequências da sua utilização, nomeadamente a possível ameaça aos



direitos civis e eventuais interpretações distorcidas das diferenças genéticas entre populações humanas que poderão, como no passado, exacerbar ideais racistas. Também a utilização inapropriada desse conhecimento, nomeadamente ao nível da seleção de embriões com base na aparência prevista através da análise de DNA é outro argumento na mesa de discussão.

Em favor da utilização destes métodos no âmbito da investigação forense são apontados argumentos como a poupança de tempo e recursos essenciais à investigação e as vantagens da previsão com base em dados concretos, estatisticamente suportados comparativamente ao relato de testemunhas oculares, que podem ser influenciadas por fatores como as circunstâncias nas quais é realizada a identificação e também eventuais preconceitos étnicos ou raciais. Outras vantagens que advêm do conhecimento relativo à diversidade genética da espécie humana, nomeadamente a melhoria do planeamento dos ensaios clínicos, a identificação de genes relacionados com doenças através de mapeamento, e o conhecimento relativo à história e evolução da espécie humana.

Em Portugal, as razões apontadas para a não utilização destes métodos são a falta de meios técnicos e restrições legais. Em relação aos meios técnicos, alguns dos painéis descritos, quer de inferência da ancestralidade geográfica (painel SNPforID 34-plex (PHILLIPS et al., 2007)), quer de previsão direta (o sistema de previsão da cor dos olhos IrisPlex (WALSH et al., 2011) e o painel de previsão da cor do cabelo descrito por (BRANICKI et al., 2011)) não são muito exigentes sob o ponto de vista de equipamentos necessários (o método de genotipagem utilizado é a minissequenciação, realizada em aparelhos de sequenciação por eletroforese capilar vulgarmente usados na genotipagem de STRs na identificação) ou até quantidade e qualidade de material exigido, visto terem sido construídos e otimizados com a finalidade de serem aplicados a amostras forenses. Sob o ponto de vista legal, a legislação regulamenta apenas, no âmbito da análise forense de DNA, a utilização de marcadores de identificação. No entanto, as restrições descritas no que toca à informação que pode ser obtida a partir de amostras de DNA (marcadores não codificantes), juntamente com o facto de a legislação ser muito protetora principalmente em aspetos relacionados com a privacidade e o facto de o DNA ser considerado um dado pessoal tornam remotas as perspetivas de aplicação da previsão de características físicas através da análise do DNA no âmbito forense, pelo menos até que haja alguma alteração à legislação ou criação de legislação específica.

A aplicação dos métodos de previsão apresenta especial interesse em Portugal tendo em conta o facto de que a base de dados de DNA é bastante recente e pouco abrangente devido ao lento crescimento do número de perfis que constam na mesma como consequência das restrições legais relativas aos procedimentos de inclusão. Este cenário torna previsível que a utilidade da base de dados seja actualmente muito limitada, tornando ainda mais importante o desenvolvimento de outros métodos que auxiliem a investigação. É importante, no entanto, enfatizar o facto de que os marcadores de previsão não substituem os marcadores de identificação em nenhuma circunstância, constituindo somente uma forma de extrair informação relativa à aparência do dador de uma amostra biológica que contenha DNA de forma a restringir o número de suspeitos a considerar nas investigações com base nessas mesmas características. A posterior utilização de marcadores de identificação continuará a ser essencial para que seja possível determinar que a amostra pertence de facto a um determinado indivíduo, ou seja, identificar o dador da amostra.

As metodologias atualmente em utilização no campo da análise forense de DNA obedecem a critérios muito rigorosos de qualidade permitindo a obtenção de dados a partir de amostras de reduzida quantidade e muitas vezes degradadas. Apresentam, no entanto, algumas limitações particularmente no que diz respeito à quantidade de marcadores que é possível genotipar em cada análise, o que constitui uma desvantagem relevante se considerarmos a hipótese de implementar a utilização de painéis de AIMs e de previsão de características físicas. Perspetiva-se que essa limitação seja ultrapassada num futuro próximo com o desenvolvimento e implementação das plataformas de sequenciação de nova geração (NGS), que, atualmente, ainda não cumprem os critérios de qualidade exigidos na análise forense.

Outro aspeto que importa referir prende-se com o facto de que os estudos de associação (usados para identificar as bases genéticas das características fenotípicas) apenas em consideração a relação entre variação fenotípica e variação genotípica ao nível da sequência de DNA. Sabe-se, no entanto, que alterações não sequenciais (epigenéticas) como, por exemplo, a ocorrência de metilações, influencia a forma como o DNA é expresso e, inevitavelmente, o fenótipo observado, para além de serem, também elas, transmitidas à descendência (KIM et al., 2009). Existem também eventos epigenéticos aleatórios, estando descrita a relação entre a ocorrência dos mesmos ao nível da morfogénese da íris e a variação observada em termos de padrão do tecido por (DAUGMAN et al., 2001). No

entanto, os estudos revistos no âmbito da realização deste trabalho não têm em consideração a potencial influência desses fenômenos.



## REFERÊNCIAS BIBLIOGRÁFICAS

ADAMS, S. M. [et al.] - The Genetic Legacy of Religious Diversity and Intolerance: Paternal Lineages of Christians, Jews, and Muslims in the Iberian Peninsula. The American Journal of Human Genetics. ISSN 0002-9297. Vol. 83, n.º 6 (2008), p. 725-736.

AHMADIAN, A.; EHN, M.; HOBER, S. - Pyrosequencing: History, biochemistry and future. Clinica Chimica Acta. ISSN 0009-8981. Vol. 363, n.º 1-2 (2006), p. 83-94.

ALPER, J. S. - Genes, free will, and criminal responsibility. Social Science & Medicine. ISSN 0277-9536. Vol. 46, n.º 12 (1998), p. 1599-1611.

AMORIM, A.; GUSMÃO, L.; ALVES, C. - STR data (AmpFI STR Profiler Plus) from north Portugal. Forensic Science International. ISSN 0379-0738. Vol. 115, n.º 1-2 (2001), p. 119-121.

ANDRÉS, A. M. [et al.] - Balancing Selection Maintains a Form of *ERAP2* that Undergoes Nonsense-Mediated Decay and Affects Antigen Presentation. PLoS Genet. Vol. 6, n.º 10 (2010), p. e1001157.

ARGAC, D. [et al.] - Optimization and validation of 10 mitochondrial DNA SNPs using SNaPshot Kit. Forensic Science International: Genetics Supplement Series. ISSN 1875-1768. Vol. 2, n.º 1 (2009), p. 99-101.

AULCHENKO, Y. S. [et al.] - Predicting human height by Victorian and genomic methods. Eur J Hum Genet. ISSN 1018-4813. Vol. 17, n.º 8 (2009), p. 1070-1075.

BAMSHAD, M. J. [et al.] - Human Population Genetic Structure and Inference of Group Membership. American Journal of Human Genetics. ISSN 0002-9297. Vol. 72, n.º 3 (2003), p. 578-589.

BAMSHAD, M. [et al.] - Deconstructing the relationship between genetics and race. Nat Rev Genet. ISSN 1471-0056. Vol. 5, n.º 8 (2004), p. 598-609.

BANDELT, H.; SALAS, A. - Current Next Generation Sequencing technology may not meet forensic standards. Forensic Science International: Genetics. ISSN 1872-4973. n.º In Press, Corrected Proof (2011).

BATES, G.; LEHRACH, H. - Trinucleotide repeat expansions and human genetic disease. BioEssays. ISSN 1521-1878. Vol. 16, n.º 4 (1994), p. 277-284.

## Referências bibliográficas

BAUCHET, M. [et al.] - Measuring European Population Stratification with Microarray Genotype Data. American Journal of Human Genetics. ISSN 0002-9297. Vol. 80, n.º 5 (2007), p. 948-956.

BBC NEWS - Doctor trials laser treatment to change eye colour [em linha]. BBC News. [Consult. 08-11-2011]Disponível em WWW:<URL:http://www.bbc.co.uk/news/technology-15596885>.

BEJERANO, G. [et al.] - Ultraconserved Elements in the Human Genome. Science. Vol. 304, n.º 5675 (2004), p. 1321-1325.

BELEZA, S. [et al.] - Micro-Phylogeographic and Demographic History of Portuguese Male Lineages. Annals of Human Genetics. ISSN 1469-1809. Vol. 70, n.º 2 (2006), p. 181-194.

BIANCHI, D. W. - Circulating Fetal DNA: Its Origin and Diagnostic Potential--A Review. Placenta. ISSN 0143-4004. Vol. 25, n.º Supplement 1 (2004), p. S93-S101.

BIO-RAD LABORATORIES, INC- Real-Time PCR Application Guide. 2006.

BIRREN, B. [et al.] - Genome Analysis – A Laboratory Manual. Cold Spring Harbor Laboratory Press, 1997.

BOEHRINGER, S. [et al.] - Genetic determination of human facial morphology: links between cleft-lips and normal variation. Eur J Hum Genet. ISSN 1018-4813. Vol. 19, n.º 11 (2011), p. 1192-1197.

BÖLL, I.; EBENBICHLER, C.; ANKENBAUER, W. - Establishing Multiplex PCR Assays Using the New LightCycler Multiplex DNA Master HybProbes and LightCycler 2.0 Instrument. Biochemica. n.º 2 (2004), p. 4-5.

BONDY, B.; BUETTNER, A.; ZILL, P. - Genetics of suicide. Mol Psychiatry. ISSN 1359-4184. Vol. 11, n.º 4 (2006), p. 336-351.

BONILLA, C. [et al.] - Ancestral proportions and their association with skin pigmentation and bone mineral density in Puerto Rican women from New York city. Human Genetics. ISSN 0340-6717. Vol. 115, n.º 1 (2004), p. 57-68.

BRANCO, C. C. [et al.] - Assessment of Azorean ancestry by Alu insertion polymorphisms. American Journal of Human Biology. ISSN 1520-6300. Vol. 18, n.º 2 (2006), p. 223-226.

BRANICKI, W. [et al.] - Model-based prediction of human hair color using DNA variants. Hum Genet. ISSN 0340-6717. (2011).

BREHM, A. [et al.] - Mitochondrial portraits of the Madeira and Açores archipelagos witness different genetic pools of its settlers. Human Genetics. ISSN 0340-6717. Vol. 114, n.º 1 (2003), p. 77-86.

BRINKMANN, B. [et al.] - Mutation Rate in Human Microsatellites: Influence of the Structure and Length of the Tandem Repeat. American Journal of Human Genetics. Vol. 62 (1998), p. 1408-1415.

BROWN, T. A. - Genomes. 2ª edição. Oxford: BIOS Science, 2002.

- BUDOWLE, B.; VAN DAAL, A. - Forensically relevant SNP classes. BioTechniques. Vol. 44, n.º 5 (2008), p. 603-610.
- BUDOWLE, B.; VAN DALL, A. - Extracting evidence from forensic DNA analyses: future molecular biology directions. BioTechniques. Vol. 46, n.º 5 (2009), p. 339-350.
- BUHLER, S.; SANCHEZ-MAZAS, A. - HLA DNA Sequence Variation among Human Populations: Molecular Signatures of Demographic and Selective Events. PLoS ONE. Vol. 6, n.º 2 (2011), p. e14643.
- BULBUL, O. [et al.] - Internal validation of 29 autosomal SNP multiplex using a ABI 310 genetic analyser. Forensic Science International: Genetics Supplement Series. ISSN 1875-1768. Vol. 2, n.º 1 (2009), p. 129-130.
- BUTLER, J.M. - Fundamentals of Forensic DNA Typing. Academic Press/Elsevier, 2009. ISBN 9780123749994.
- BUTLER, J. M.; REEDER, D. J. - Short Tandem Repeat Internet DataBase [em linha]. [Consult. 06-04-2010]. Disponível em WWW:<URL:<http://www.cstl.nist.gov/strbase/index.htm>>.
- CADENAS, A. M. [et al.] - Male amelogenin dropouts: phylogenetic context, origins and implications. Forensic Science International. ISSN 0379-0738. Vol. 166, n.º 2-3 (2007), p. 155-163.
- CALAFELL, F.; COMAS, D.; BERTRANPETIT, J. - Why Names. Genome Research. Vol. 12, n.º 2 (2002), p. 219-221.
- CALDARA, R.; ZHOU, X.; MIELLET, S. - Putting Culture Under the 'Spotlight' Reveals Universal Information Use for Face Recognition. PLoS ONE. Vol. 5, n.º 3 (2010), p. e9708.
- CARVALHO, M. [et al.] - Y chromosome SNP analysis in a Portuguese population sample using a multiplex PCR and minisequencing strategy. Forensic Science International: Genetics Supplement Series. ISSN 1875-1768. Vol. 1, n.º 1 (2008), p. 190-191.
- CAVALLI-SFORZA, L. L. - The Human Genome Diversity Project: past, present and future. Nat Rev Genet. ISSN 1471-0056. Vol. 6, n.º 4 (2005), p. 333-340.
- CAVALLI-SFORZA, L. L.; FELDMAN, M. W. - The application of molecular genetic approaches to the study of human evolution. Nature Genetics Supplement. Vol. 33 (2003), p. 266-275.
- CAVALLI-SFORZA, L. L.; MENOZZI, P.; PIAZZA, A. - The History and Genography of Human Genes. Princeton: Princeton University Press, 1994.
- CHEN, T. [et al.] - The mitochondrial DNA 4,977-bp deletion and its implication in copy number alteration in colorectal cancer. BMC Medical Genetics. ISSN 1471-2350. Vol. 12, n.º 1 (2011), p. 8.
- COLLINS-SCHRAMM, H. [et al.] - Mexican American ancestry-informative markers: examination of population structure and marker characteristics in European Americans,

## Referências bibliográficas

Mexican Americans, Amerindians and Asians. Human Genetics. ISSN 0340-6717. Vol. 114, n.º 3 (2004), p. 263-271.

COLLINS, F. S.; BROOKS, L. D.; CHAKRAVARTI, A. - A DNA polymorphism discovery resource for research on human genetic variation. Genome Research. Vol. 8 (1998), p. 1229-1231.

CONE, A. J. - DNA and Genealogy [em linha]. [Consult. 28-02-2011] Disponível em WWW:<URL:http://www.dnainheritance.kahikatea.net/>.

CONRAD, D. F.; HURLES, M. E. - The population genetics of structural variation. Nat Genet. (2007).

CONSELHO NACIONAL DE ÉTICA PARA AS CIÊNCIAS DA VIDA - PARECER SOBRE O REGIME JURÍDICO DA BASE DE DADOS DE PERFIS DE A.D.N. [em linha]. (2007). [Consult. 05-09-2011] Disponível na internet:<URL:http://www.cneqv.pt/admin/files/data/docs/1273054082\_Parecer\_052\_CNECV\_2007\_BasesdadosADN.pdf>.

CONSELHO NACIONAL DE ÉTICA PARA AS CIÊNCIAS DA VIDA - RELATÓRIO E PARECER SOBRE IMPLICAÇÕES ÉTICAS DA GENÓMICA. [em linha]. (2001). [Consult. 05-09-2011] Disponível na internet:<URL:http://www.cneqv.pt/admin/files/data/docs/1273057350\_P040\_ImplicacoesEticasGenomica.pdf>.

COSTA, G. [et al.] - Autosomal SNPs in paternity investigation. Forensic Science International: Genetics Supplement Series. ISSN 1875-1768. Vol. 1, n.º 1 (2008), p. 507-509.

CRAIG, I.; HALTON, K. - Genetics of human aggressive behaviour. Human Genetics. ISSN 0340-6717. Vol. 126, n.º 1 (2009), p. 101-113.

DALE, J. W.; VON SCHANTZ, M. - From genes and genomes: concepts and applications of DNA technology. 2ª edição. West Sussex: Wiley, 2002.

DAUBER, E. M. [et al.] - Discrepant results of samples taken from different tissues of a single individual. International Congress Series. ISSN 0531-5131. Vol. 1261 (2004), p. 48-49.

DAUGMAN, J.; DOWNING, C. - Epigenetic randomness, complexity and singularity of human iris patterns. Proceeding of the Royal Society (London): B. Biological Sciences. Vol. 268 (2001), p. 1737-1740.

DNA PRINT GENOMICS - DNAPrint Announces The Release Of RETINOME™ For The Forensic Market: Eye Color Prediction From Crime Scene DNA [em linha]. [Consult. 26-04-2011] Disponível em WWW:<URL:http://www.dnaprint.com/welcome/press/press\_recent/2004/august\_17/>.

DNA PRINT GENOMICS - DNAPRINT GENOMICS ANNOUNCES FIRST SCHEDULED SALES OF UPGRADED DNAWITNESS PRODUCT FEATURING RETINOME 2.0 [em linha]. [Consult. 26-04-2011] Disponível em



WWW:<URL:[http://www.dnaprint.com/welcome/press/press\\_recent/2006/0810/DNAG-RETINOME2.pdf](http://www.dnaprint.com/welcome/press/press_recent/2006/0810/DNAG-RETINOME2.pdf)>.

DNA PRINT GENOMICS - Products and Services [em linha]. Sarasota. [Consult. 15-04-2011]. Disponível em WWW:<URL:<http://www.dnaprint.com/welcome/home/>>.

DUFFY, D. L. [et al.] - A three-single-nucleotide polymorphism haplotype in intron 1 of OCA2 explains most human eye-color variation. The American Journal of Human Genetics. Vol. 80 (2007), p. 241-252.

DUFFY, D. L. [et al.] - Multiple Pigmentation Gene Polymorphisms Account for a Substantial Proportion of Risk of Cutaneous Malignant Melanoma. J Invest Dermatol. ISSN 0022-202X. Vol. 130, n.º 2 (2009), p. 520-528.

EIBERG, H. [et al.] - Blue eye color in humans may be caused by a perfectly associated founder mutation in a regulatory element located within the HERC2 gene inhibiting OCA2 expression. Human Genetics. ISSN 0340-6717. Vol. 123, n.º 2 (2008), p. 177-187.

ELLEGREN, H. - Microsatellites: simple sequences with complex evolution. Nature Reviews Genetics. Vol. 5 (2004), p. 435-445.

ERIKSSON, N. [et al.] - Web-Based, Participant-Driven Studies Yield Novel Genetic Associations for Common Traits. PLoS Genet. Vol. 6, n.º 6 (2010), p. e1000993.

EUROPEAN BIOINFORMATIC INSTITUTE - IMGT/HLA Database [em linha]. European Bioinformatic Institute. [Consult. 03-10-2011] Disponível em WWW:<URL:<http://www.ebi.ac.uk/imgt/hla/>>.

FERNANDES, A.; BREHM, A. - Y-chromosome STR haplotypes in the Açores Archipelago (Portugal). Forensic Science International. ISSN 0379-0738. Vol. 135, n.º 3 (2003), p. 239-242.

FERNANDES, A. T.; BREHM, A. - Population data of five STRs in three regions from Portugal. Forensic Science International. ISSN 0379-0738. Vol. 129, n.º 1 (2002a), p. 72-74.

FERNANDES, ANA TERESA [et al.] - Genetic profile of the Madeira Archipelago population using the new PowerPlex(R) 16 System kit. Forensic Science International. Vol. 125 (2002b), p. 281-283.

FERNANDES, A. T. [et al.] - Y-chromosome STR haplotypes in the Madeira archipelago population. Forensic Science International. ISSN 0379-0738. Vol. 122, n.º 2-3 (2001), p. 178-180.

FLORES, C. [et al.] - Reduced genetic structure of the Iberian peninsula revealed by Y-chromosome analysis: implications for population demography. Eur J Hum Genet. ISSN 1018-4813. Vol. 12, n.º 10 (2004), p. 855-863.

FOX, D. - The Second Generation of Racial Profiling. Student Prize Papers. Vol. Paper 51 (2010).

FROST, P. - European hair and eye color: A case of frequency-dependent sexual selection? Evolution and Human Behavior. ISSN 1090-5138. Vol. 27, n.º 2 (2006), p. 85-103.

FRUDAKIS, T. [et al.] - A Classifier for the SNP-Based Inference of Ancestry. Journal Forensic Sciences. Vol. 48, n.º 4 (2003), p. 771-82.

FRUDAKIS, T. N. - Molecular photofitting: Predicting Ancestry and Phenotype Using DNA. Academic Press Publishers (Elsevier), 2008.

FUJIMOTO, A. [et al.] - A scan for genetic determinants of human hair morphology: EDAR is associated with Asian hair thickness. Hum. Mol. Genet.. Vol. 17 (2008), p. 835-843.

FUJIMOTO, A. [et al.] - FGFR2 is associated with hair thickness in Asian populations. . Journal of Human Genetics. Vol. 54 (2009), p. 461-465.

GEADA, H. [et al.] - Portuguese population and paternity investigation studies with a multiplex PCR -- the AmpFI STR® Profiler Plus(TM). Forensic Science International. ISSN 0379-0738. Vol. 108, n.º 1 (2000), p. 31-37.

GENOMEWEB DAILY NEWS - DNAPrint Genomics Goes Bust [em linha]. [Consult. 15-04-2011] Disponível em WWW:<URL:http://www.genomeweb.com/dnaprint-genomics-goes-bust>.

GILL, P. [et al.] - An assessment of whether SNPs will replace STRs in national DNA databases--joint considerations of the DNA working group of the European Network of Forensic Science Institutes (ENFSI) and the Scientific Working Group on DNA Analysis Methods (SWGDM). Science & justice : journal of the Forensic Science Society. Vol. 44, n.º 1 (2004), p. 51-3.

GOLDSBY, R. A.; KINDT, T. J.; OSBORNE, B. A. - Kuby Immunology. 5. W. B. Saunders Company, 2004.

GONÇALVES, R. [et al.] - Y-chromosome lineages from Portugal, Madeira and Açores record elements of Sephardim and Berber ancestry. International Congress Series. ISSN 0531-5131. Vol. 1288 (2006), p. 237-239.

GONZÁLEZ-NEIRA, A. [et al.] - Distribution of Y-chromosome STR defined haplotypes in Iberia. Forensic Science International. ISSN 0379-0738. Vol. 110, n.º 2 (2000), p. 117-126.

GONZALEZ, KELLY D. [et al.] - Somatic microindels: analysis in mouse soma and comparison with the human germline. Human Mutation. ISSN 1098-1004. Vol. 28, n.º 1 (2007), p. 69-80.

GRAHAM, E. A. M. - DNA reviews: predicting phenotype. Forensic Science Medical Pathology. Vol. 4 (2008), p. 196-199.

GRIMES, E. A. [et al.] - Sequence polymorphism in the human melanocortin 1 receptor gene as an indicator of the red hair phenotype. Forensic Science International. ISSN 0379-0738. Vol. 122, n.º 2 (2001), p. 124-129.

GUIMARÃES, A. S. A. - Racismo e anti-racismo no Brasil. FUSP, Fundação de Apoio à Universidade de São Paulo, 2005. ISBN 9788573261394.

GUSMÃO, A. [et al.] - A Perspective on the History of the Iberian Gypsies Provided by Phylogeographic Analysis of Y-Chromosome Lineages. Annals of Human Genetics. ISSN 1469-1809. Vol. 72, n.º 2 (2008a), p. 215-227.

GUSMÃO, A. [et al.] - Y-chromosomal STR haplotypes in a Gypsy population from Portugal. Forensic Science International: Genetics Supplement Series. ISSN 1875-1768. Vol. 1, n.º 1 (2008b), p. 212-213.

GUSMÃO, A. [et al.] - A genetic historical sketch of European Gypsies: The perspective from autosomal markers. American Journal of Physical Anthropology. ISSN 1096-8644. Vol. 141, n.º 4 (2010), p. 507-514.

HAMILTON, G.; STONEKING, M.; EXCOFFIER, L. - Molecular analysis reveals tighter social regulation of immigration in patrilocal populations than in matrilocal populations. PNAS. Vol. 102, n.º 21 (2005), p. 7476-7480.

HAN, J. [et al.] - A genome-wide association study identifies novel alleles associated with hair color and skin pigmentation. PLoS Genetics. Vol. 4, n.º 5 (2008), p. e1000074.

HANCOCK, A. M. [et al.] - Adaptations to Climate-Mediated Selective Pressures in Humans. PloS Genetics. Vol. 7, n.º 4 (2011), p. e1001375.

HARRISON, C. [et al.] - A sensitive issue: Pyrosequencing as a valuable forensic SNP typing platform. International Congress Series. ISSN 0531-5131. Vol. 1288 (2006), p. 52-54.

HE, Y. [et al.] - Heteroplasmic mitochondrial DNA mutations in normal and tumour cells. Nature. ISSN 0028-0836. Vol. 464, n.º 7288 (2010), p. 610-614.

HIRSCHHORN, J. N.; LETTRE, G. - Progress in Genome-Wide Association Studies of Human Height. Hormone Research. Vol. 71, n.º Suppl. 2 (2009), p. 5-13.

HOLSINGER, K. E.; WEIR, B. S. - Genetics in geographically structured populations: defining, estimating and interpreting FST. Nat Rev Genet. ISSN 1471-0056. Vol. 10, n.º 9 (2009), p. 639-650.

- Method for Alteration of Iris Pigmentation. HOMER, GREGG S. Estados Unidos da América

HUDSON, R. R.; SLATKIN, M.; MADDISON, W. P. - Estimation of Levels of Gene Flow From DNA Sequence Data. Genetics. Vol. 132, n.º 2 (1992), p. 583-589.

IL'INA, E.; GOVORUN, V. - Mass spectrometry of nucleic acids in molecular medicine. Russian Journal of Bioorganic Chemistry. ISSN 1068-1620. Vol. 35, n.º 2 (2009), p. 135-149.

Infopédia - Enciclopédia e Dicionários Porto Editora - [em linha]. Porto Editora. [Consult. 05-09-2011]. Disponível em WWW:<URL:http://www.infopedia.pt/>.

INNOCENCE PROJECT - 250 Exonerated, Too Many Wrongfully Convicted. [em linha]. (2010). [Consult. 14-06-2011] Disponível na internet:<URL:http://www.innocenceproject.org/news/250.php>.

INTERNATIONAL HUMAN GENOME SEQUENCING CONSORTIUM - Initial sequencing and analysis of the human genome. Nature. ISSN 0028-0836. Vol. 409, n.º 6822 (2001), p. 860-921.

IRWIN, J. A. [et al.] - mtGenome reference population databases and the future of forensic mtDNA analysis. Forensic Science International: Genetics. ISSN 1872-4973. Vol. 5, n.º 3 (2011), p. 222-225.

JABLONSKI, N. G.; CHAPLIN, G. - Human skin pigmentation as an adaptation to UV radiation. Proceedings of the National Academy of Sciences. Vol. 107, n.º Supplement 2 (2010), p. 8962-8968.

JACKSON JR, J. P. - "In Ways Unacademical": The Reception of Carleton S. Coon's The Origin of Races. Journal of the History of Biology. Vol. 34, n.º 2 (2001), p. 247-285.

JOBLING, M. - Elementary, my dear Cameron. Investigative Genetics. ISSN 2041-2223. Vol. 2, n.º 1 (2011), p. 5.

JOBLING, M. A.; GILL, P. - Encoded evidence: DNA in forensic analysis. Nature Reviews Genetics. Vol. 5 (2004), p. 739-751.

JOBLING, M. A.; TYLER-SMITH, C. - The human Y chromosome: an evolutionary marker comes of age. Nat Rev Genet. ISSN 1471-0056. Vol. 4, n.º 8 (2003), p. 598-612.

JOHANSON, D. - Origins of Modern Humans: Multiregional or Out of Africa? [em linha]. ActionBioscience.org. [Consult. 12-01-2011] Disponível em WWW:<URL:<http://www.actionbioscience.org/evolution/johanson.html>>.

JORDE, L. B.; WATKINS, W. S.; BAMSHAD, M. J. - Population genomics: a bridge from evolutionary history to genetic medicine. Human Molecular Genetics. Vol. 10, n.º 20 (2001), p. 2199-2207.

JORDE, LYNN B.; WOODING, STEPHEN P. - Genetic variation, classification and 'race'. Nat Genet. (2004).

JUNQUEIRA, L. C.; CARNEIRO, J. - Histologia Básica. Guanabara Koogan, 2004.

JURMAIN, R.; KILGORE, L.; TREVATHAN, W. - Essentials of Physical Anthropology. Wadsworth, 2008. ISBN 9780495509394.

KASHYAP, V. K. [et al.] - Deletions in the Y-derived amelogenin gene fragment in the Indian population. BMC Medical Genetics. ISSN 1471-2350. Vol. 7, n.º 1 (2006), p. 37.

KAYSER, M. - Uni-parental markers in human identity testing including forensic DNA analysis. BioTechniques. Vol. 43, n.º 6 (2007), p. Sxxv-Sxxi.

KAYSER, M. [et al.] - Three genome-wide association studies and linkage analysis identify HERC2 as a human iris color gene. The American Journal of Human Genetics. Vol. 82 (2008), p. 411-423.

- KAYSER, M.; SCHNEIDER, P. M. - DNA-based prediction of human externally visible characteristics in forensics: Motivations, scientific challenges, and ethical considerations. Forensic Science International: Genetics. Vol. 3 (2009), p. 154-161.
- KIDD, J. [et al.] - Analyses of a set of 128 ancestry informative single-nucleotide polymorphisms in a global set of 119 population samples. Investigative Genetics. ISSN 2041-2223. Vol. 2, n.º 1 (2011), p. 1.
- KIM, J.; SAMARANAYAKE, M.; PRADHAN, S. - Epigenetic mechanisms in mammals. Cellular and Molecular Life Sciences. ISSN 1420-682X. Vol. 66, n.º 4 (2009), p. 596-612.
- KOHN, L. A. P. - THE ROLE OF GENETICS IN CRANIOFACIAL MORPHOLOGY AND GROWTH. Annual Review of Anthropology. ISSN 0084-6570. Vol. 20 (1991), p. 261-278.
- KOOPS, B.; SCHELLEKENS, M. - Forensic DNA Phenotyping: Regulatory Issues. The Columbia Science and Technology Law Review. Vol. 9 (2008), p. 158-202.
- KWOK, P. Y. - Methods for genotyping single nucleotide polymorphisms. Annual Review of Genomics and Human Genetics. Vol. 2 (2001), p. 235-58.
- LAGOA, A. M. [et al.] - Allele frequencies of six miniSTR loci in the population of Northern Portugal. Forensic Science International: Genetics. ISSN 1872-4973. Vol. 2, n.º 4 (2008), p. 379-381.
- LANGO ALLEN, H. [et al.] - Hundreds of variants clustered in genomic loci and biological pathways affect human height. Nature. ISSN 0028-0836. Vol. 467, n.º 7317 (2010), p. 832-838.
- LAO, O. [et al.] - Signatures of Positive Selection in Genes Associated with Human Skin Pigmentation as Revealed from Analyses of Single Nucleotide Polymorphisms. Annals of Human Genetics. ISSN 1469-1809. Vol. 71, n.º 3 (2007), p. 354-369.
- LAO, O. [et al.] - Proportioning Whole-Genome Single-Nucleotide-Polymorphism Diversity for Identification of Geographic Population Structure and Genetic Ancestry. The American Journal of Human Genetics. Vol. 78 (2006), p. 680-690.
- LAREU, M. [et al.] - The use of the LightCycler for the detection of Y chromosome SNPs. Forensic Science International. ISSN 0379-0738. Vol. 118, n.º 2-3 (2001), p. 163-168.
- LEBRECHT, S. [et al.] - Perceptual Other-Race Training Reduces Implicit Racial Bias. PLoS ONE. Vol. 4, n.º 1 (2009), p. e4215.
- LEE, S. [et al.] - The ethics of characterizing difference: guiding principles on using racial categories in human genetics. Genome Biology. ISSN 1465-6906. Vol. 9, n.º 7 (2008), p. 404.
- LEI, S. [et al.] - Genome-wide association scan for stature in Chinese: evidence for ethnic specific loci. Human Genetics. ISSN 0340-6717. Vol. 125, n.º 1 (2009), p. 1-9.
- LENZ, W. - Phenocopy. Human Genetics. ISSN 0340-6717. Vol. 9, n.º 3 (1970), p. 227-229.

LESSIG, R. [et al.] - SNP-genotyping on human Y-chromosome for forensic purposes: comparison of two different methods. International Congress Series. ISSN 0531-5131. Vol. 1261 (2004), p. 334-336.

LETTRE, G. - Genetic regulation of adult stature. Current Opinion in Pediatrics. Vol. 21, n.º 4 (2009), p. 515-22.

LEWIN, R. - Molecular Anthropology. In: Patterns in Evolution: The New Molecular View. New York: Scientific American Library, 1997.

LI, J. Z. [et al.] - Worldwide Human Relationships Inferred from Genome-Wide Patterns of Variation. Science. Vol. 319, n.º 5866 (2008), p. 1100-1104.

LI, M. [et al.] - Detecting Heteroplasmy from High-Throughput Sequencing of Complete Human Mitochondrial DNA Genomes. American Journal of Human Genetics. Vol. 87, n.º 2 (2010), p. 237-249.

LIMA, G. [et al.] - Analysis of the HVI and HVII regions of mitochondrial DNA in 100 individuals from North of Portugal. International Congress Series. ISSN 0531-5131. Vol. 1261 (2004), p. 366-368.

LIMER, E. - There Is A Laser Surgery That Can Permanently Change Eye Color [em linha]. Geek O System. [Consult. 08-11-2011] Disponível em WWW:<URL:<http://www.geekosystem.com/laser-surgery-eye-color/>>.

LINDSAY, D. S.; JACK, P. C.; CHRISTIAN, M. A. - Other-race face perception. Journal of Applied Psychology. Vol. 76, n.º 4 (1991), p. 587-589.

LIU, F. [et al.] - Eye color and the prediction of complex phenotypes from genotypes. Current biology : CB. ISSN 0960-9822. Vol. 19, n.º 5 (2009), p. R192-R193.

LIU, F. [et al.] - Digital Quantification of Human Eye Color Highlights Genetic Association of Three New Loci. PLoS Genet. Vol. 6, n.º 5 (2010), p. e1000934.

LIVAK, K. J. - Allelic discrimination using fluorogenic probes and the 5' nuclease assay. Genetic Analysis: Biomolecular Engineering. ISSN 1050-3862. Vol. 14, n.º 5-6 (1999), p. 143-149.

LOUSSOUARN, G. [et al.] - Worldwide diversity of hair curliness: a new method of assessment. International Journal of Dermatology. Vol. 46, n.º Suppl 1 (2007), p. 2-6.

MANNUCCI, A. [et al.] - Forensic application of a rapid and quantitative DNA sex test by amplification of the X-Y homologous gene amelogenin. International Journal of Legal Medicine. ISSN 0937-9827. Vol. 106, n.º 4 (1994), p. 190-193.

MAO, X. [et al.] - A Genomewide Admixture Mapping Panel for Hispanic/Latino Populations. American Journal of Human Genetics. Vol. 80, n.º 6 (2007), p. 1171-1178.

MARDIS, E. R. - Next-generation DNA sequencing methods. In: Annual Review of Genomics and Human Genetics. Palo Alto: Annual Reviews, 2008. ISBN 1527-8204 978-0-8243-3709-4, p. 387-402.

- MCCARTHY, M. I. [et al.] - Genome-wide association studies for complex traits: consensus, uncertainty and challenges. Nat Rev Genet. ISSN 1471-0056. Vol. 9, n.º 5 (2008), p. 356-369.
- MEDLAND, S. E. [et al.] - Common Variants in the Trichohyalin Gene Are Associated with Straight Hair in Europeans. American Journal of Human Genetics. ISSN 0002-9297. Vol. 85, n.º 5 (2009), p. 750-755.
- MEISSNER, C.; RITZ-TIMME, S. - Molecular pathology and age estimation. Forensic Science International. ISSN 0379-0738. Vol. 203, n.º 1-3 (2010), p. 34-43.
- MENGEL-FROM, J. [et al.] - Human eye colour and HERC2, OCA2 and MATP. Forensic Science International: Genetics. ISSN 1872-4973. Vol. 4, n.º 5 (2010), p. 323-328.
- MESSIER, W.; LI, S.; STEWART, C. - The birth of microsatellites. Nature. Vol. 381, n.º 6582 (1996), p. 483-483.
- METZKER, M. L. - Sequencing technologies - the next generation. Nat Rev Genet. ISSN 1471-0056. Vol. 11, n.º 1 (2010), p. 31-46.
- MICHAEL, A.; BRAUNER, P. - Erroneous Gender Identification by the Amelogenin Sex Test. Journal Forensic Sciences. Vol. 49, n.º 2 (2004).
- MONTIEL, R. [et al.] - Analysis of Y-chromosome Variability and its Comparison with mtDNA Variability Reveals Different Demographic Histories Between Islands in the Azores Archipelago (Portugal). Annals of Human Genetics. ISSN 1469-1809. Vol. 69, n.º 2 (2005), p. 135-144.
- MOSQUERA-MIGUEL, A. [et al.] - Testing the performance of mtSNP minisequencing in forensic samples. Forensic Science International: Genetics. ISSN 1872-4973. Vol. 3, n.º 4 (2009), p. 261-264.
- MULLANEY, JULIENNE M. [et al.] - Small insertions and deletions (INDELs) in human genomes. Human Molecular Genetics. Vol. 19, n.º R2 (2010), p. R131-R136.
- NATIONAL FORENSIC SCIENCE TECHNOLOGY CENTER (NFSTC) - President's DNA Initiative - DNA Analysts training website [em linha]. Largo, Florida: WorldPress. [Consult. 28-02-2011] Disponível em WWW:<URL:<http://www.nfstc.org/>>.
- NATIONAL GEOGRAPHIC SOCIETY - National Geographic - The Genographic Project [em linha]. National Geographic Society. [Consult. 27-01-2011] Disponível em WWW:<URL:<https://genographic.nationalgeographic.com/genographic/index.html>>.
- NIEDERSTÄTTER, H. [et al.] - Characterization of mtDNA SNP typing and mixture ratio assessment with simultaneous real-time PCR quantification of both allelic states. International Journal of Legal Medicine. ISSN 0937-9827. Vol. 120, n.º 1 (2006), p. 18-23.
- NOGUEIRO, I. [et al.] - Phylogeographic analysis of paternal lineages in NE Portuguese Jewish communities. American Journal of Physical Anthropology. ISSN 1096-8644. Vol. 141, n.º 3 (2010), p. 373-381.

- NORTON, H. L. [et al.] - Genetic Evidence for the Convergent Evolution of Light Skin in Europeans and East Asians. Molecular Biology and Evolution. Vol. 24, n.º 3 (2007), p. 710-722.
- NOVEMBRE, J. [et al.] - Genes mirror geography within Europe. Nature. Vol. 456, n.º 7218 (2008), p. 98-101.
- OBASOGIE, O. - Playing the Gene Card? Oakland, CA: Center for Genetics and Society, 2009.
- ÖZALP, E. - The Genetics of Suicidal Behavior. Turkish Journal of Psychiatry. Vol. 20, n.º 1 (2009), p. 85-93.
- PACHECO, P. R. [et al.] - The Y-chromosomal Heritage of the Azores Islands Population. Annals of Human Genetics. ISSN 1469-1809. Vol. 69, n.º 2 (2005), p. 145-156.
- PANZER, S.; KUHLE, D. P. A.; CASKEY, C. T. - Unstable triplet sequences: a source of cancer mutations? Stem Cells. Vol. 13 (1995), p. 146-157.
- PARK, P. J. - ChIP-seq: advantages and challenges of a maturing technology. Nat Rev Genet. ISSN 1471-0056. Vol. 10, n.º 10 (2009), p. 669-680.
- PARRA, E. J.; KITTLES, R. A.; SHRIVER, M. D. - Implications of correlations between skin color and genetic ancestry for biomedical research. Nature Genetics Supplement. Vol. 36, n.º 11 (2004), p. s54-s60.
- PARSON, W.; DÜR, A. - EMPOP—A forensic mtDNA database. Forensic science international. Genetics. ISSN 1872-4973. Vol. 1, n.º 2 (2007), p. 88-92.
- PATI, N. [et al.] - A comparison between SNaPshot, pyrosequencing, and biplex invader SNP genotyping methods: accuracy, cost, and throughput. Journal of Biochemical and Biophysical Methods. ISSN 0165-022X. Vol. 60, n.º 1 (2004), p. 1-12.
- PEREIRA, L. [et al.] - Clinal variation of the YAP+ Y chromosome frequencies in Western Iberia. Human Biology. Vol. 72 (2000a), p. 937-944.
- PEREIRA, L. [et al.] - African female heritage in Iberia: A reassessment of mtDNA lineage distribution in present times. Human Biology. ISSN 0018-7143. Vol. 77, n.º 2 (2005), p. 213-229.
- PEREIRA, L. [et al.] - Phylogeny of the mtDNA haplogroup U6. Analysis of the sequences observed in North Africa and Iberia. International Congress Series. ISSN 0531-5131. Vol. 1239 (2003), p. 491-493.
- PEREIRA, L.; PRATA, M. J.; AMORIM, A. - Diversity of mtDNA lineages in Portugal: not a genetic edge of European variation. Annals of Human Genetics. ISSN 1469-1809. Vol. 64, n.º 6 (2000b), p. 491-506.
- PEREIRA, L.; CUNHA, C.; AMORIM, A. - Predicting sampling saturation of mtDNA haplotypes: an application to an enlarged Portuguese database. International Journal of Legal Medicine. ISSN 0937-9827. Vol. 118, n.º 3 (2004), p. 132-136.



- PEREIRA, L. [et al.] - Analysis of the Y-chromosome and Mitochondrial DNA pools in Portugal. In: RENFREW, C. ; BOYLE, K. - Archaeogenetics: DNA and the population prehistory of Europe. Cambridge: Oxbow Books, 2000c. p. 191-195.
- PEREIRA, R. [et al.] - Genetic characterization of 52 autosomal SNPs in the Portuguese population. Forensic Science International: Genetics Supplement Series. ISSN 1875-1768. Vol. 1, n.º 1 (2008), p. 358-360.
- PEREIRA, V. [et al.] - Genetic characterization of uniparental lineages in populations from Southwest Iberia with past malaria endemicity. American Journal of Human Biology. ISSN 1520-6300. Vol. 22, n.º 5 (2010), p. 588-595.
- PEREIRA, V. [et al.] - Study of 25 X-chromosome SNPs in the Portuguese. Forensic Science International: Genetics. ISSN 1872-4973. Vol. 5, n.º 4 (2011), p. 336-338.
- PETKOVSKI, E. [et al.] - SNPs and MALDI-TOF MS: Tools for DNA Typing in Forensic Paternity Testing and Anthropology. Journal Forensic Sciences. Vol. 50, n.º 3 (2005).
- PHILLIPS, C. [et al.] - Inferring ancestral origin using a single multiplex assay of ancestry-informative marker SNPs. Forensic science international. Genetics. ISSN 1872-4973. Vol. 1, n.º 3 (2007), p. 273-280.
- PHILLIPS, C. - Online resources for SNP analysis. Molecular Biotechnology. ISSN 1073-6085. Vol. 35, n.º 1 (2007), p. 65-97.
- PHILLIPS, C. [et al.] - Ancestry analysis in the 11-M Madrid bomb attack investigation. PLoS ONE. Vol. 4, n.º 8 (2009), p. e6583.
- PHILLIPS, M. L. - Crime Scene Genetics: Transforming Forensic Science through Molecular Technologies. BioScience. Vol. 58, n.º 6 (2008), p. 484-489.
- PINHEIRO, M. F. [et al.] - Allele frequencies of sixteen STRs in the population of Northern Portugal. Forensic Science International. ISSN 0379-0738. Vol. 148, n.º 2-3 (2005), p. 221-223.
- POLLARD, K. S. [et al.] - Forces Shaping the Fastest Evolving Regions in the Human Genome. PLoS Genet. Vol. 2, n.º 10 (2006), p. e168.
- PONTES, M. L. [et al.] - Allele frequencies and population data for 17 Y-STR loci (AmpFISTR® Y-filer(TM)) in a Northern Portuguese population sample. Forensic Science International. ISSN 0379-0738. Vol. 170, n.º 1 (2007), p. 62-67.
- PULKER, H. [et al.] - Finding genes that underlie physical traits of forensic interest using genetic tools. Forensic Science International: Genetics. Vol. 1 (2007), p. 100-104.
- RACE ETHNICITY AND GENETICS WORKING GROUP - The Use of Racial, Ethnic, and Ancestral Categories in Human Genetics Research. American Journal of Human Genetics. ISSN 0002-9297. Vol. 77, n.º 4 (2005), p. 519-532.
- RAY, D. A.; WALKER, J. A.; BATZER, M. A. - Mobile element-based forensic genomics. Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis. ISSN 0027-5107. Vol. 616, n.º 1-2 (2007), p. 24-33.

REES, J. L. - Genetics of Hair and Skin Color. Annual Review of Genetics. Vol. 37, n.º 1 (2003), p. 67-90.

ROSENBERG, N. A. [et al.] - Informativeness of Genetic Markers for Inference of Ancestry. American Journal of Human Genetics. Vol. 73 (2003), p. 1402-1422.

ROSENBERG, N. A. [et al.] - Genetic Structure of Human Population. Science. Vol. 298 (2002), p. 2381-2385.

ROYAL, C. D. [et al.] - Inferring Genetic Ancestry: Opportunities, Challenges, and Implications. American Journal of Human Genetics. ISSN 0002-9297. Vol. 86, n.º 5 (2010), p. 661-673.

SANTOS, C. [et al.] - Genetic Structure and Origin of Peopling in The Azores Islands (Portugal): The View from mtDNA. Annals of Human Genetics. ISSN 1469-1809. Vol. 67, n.º 5 (2003), p. 433-456.

SANTOS, F. R.; PANDYA, A.; TYLER-SMITH, C. - Reliability of DNA-based sex tests. Nat Genet. Vol. 18, n.º 2 (1998), p. 103-103.

SCHWARTZ, M.; VISSING, J. - Paternal Inheritance of Mitochondrial DNA. New England Journal of Medicine. Vol. 347, n.º 8 (2002), p. 576-580.

SHIMOMURA, Y. [et al.] - Autosomal-Dominant Woolly Hair Resulting from Disruption of Keratin 74 (KRT74), a Potential Determinant of Human Hair Texture. American Journal of Human Genetics. ISSN 0002-9297. Vol. 86, n.º 4 (2010), p. 632-638.

SHRIVER, M. D. [et al.] - Large-scale SNP analysis reveals clustered and continuous patterns of human genetic variation. Human Genomics. Vol. 2, n.º 2 (2005), p. 81-89.

SILVA, F. [et al.] - Genetic profiling of the Azores Islands (Portugal): Data from 10 X-chromosome STRs. American Journal of Human Biology. ISSN 1520-6300. Vol. 22, n.º 2 (2010), p. 221-223.

SMITH, H. F. - The Role of Genetic Drift in Shaping Modern Human Cranial Evolution: A Test Using Microevolutionary Modeling. International Journal of Evolutionary Biology. Vol. 2011 (2011).

SOBRINO, B.; BRIÓN, M.; CARRACEDO, A. - SNPs in forensic genetics: a review on SNP typing methodologies. Forensic Science International. ISSN 0379-0738. Vol. 154, n.º 2-3 (2005), p. 181-194.

SPICHENOK, O. [et al.] - Prediction of eye and skin color in diverse populations using seven SNPs. Forensic Science International: Genetics. ISSN 1872-4973. Vol. In Press, Corrected Proof (2010).

SPINNEY, L. - Eyewitness identification: Line-ups on trial. Nature. Vol. 453 (2008), p. 442-444.

SPÍNOLA, H. [et al.] - Distribution of HLA alleles in Portugal and Cabo Verde. Relationships with the slave trade route. Annals of Human Genetics. ISSN 1469-1809. Vol. 66, n.º 4 (2002), p. 285-296.

STANDFORD UNIVERSITY - The Origin of Modern Man [em linha]. Standford. [Consult. 12-01-2011] Disponível em WWW:<URL:<http://www.stanford.edu/~harryg/protected/evolve4.htm>>.

STEINLECHNER, M. [et al.] - Rare failures in the amelogenin sex test. International Journal of Legal Medicine. ISSN 0937-9827. Vol. 116, n.º 2 (2002), p. 117-120.

STOKOWSKI, R. P. [et al.] - A Genomewide Association Study of Skin Pigmentation in a South Asian Population. American Journal of Human Genetics. ISSN 0002-9297. Vol. 81, n.º 6 (2007), p. 1119-1132.

STRACHAN, T.; READ, A.P. - Human molecular genetics. Garland Science, 2010. ISBN 9780815341499.

STURM, R. A. - Molecular genetics of human pigmentation diversity. Human Molecular Genetics. Vol. 18, n.º R1 (2009), p. R9-R17.

STURM, R. A.; FRUDAKIS, T. N. - Eye colour: portals into pigmentation genes and ancestry. Trends in Genetics. Vol. 20, n.º 8 (2004), p. 327-332.

SULEM, P. [et al.] - Two newly identified genetic determinants of pigmentation in Europeans. Nat Genet. ISSN 1061-4036. Vol. 40, n.º 7 (2008), p. 835-837.

SULEM, P. [et al.] - Genetic determinants of hair, eye and skin pigmentation in Europeans. Nat Genet. ISSN 1061-4036. Vol. 39, n.º 12 (2007), p. 1443-1452.

SZIBOR, R. [et al.] - Use of X-linked markers for forensic purposes. International Journal of Legal Medicine. Vol. 117 (2003), p. 67-74.

SZIBOR, R. - X-chromosomal markers: Past, present and future. Forensic science international. Genetics. ISSN 1872-4973. Vol. 1, n.º 2 (2007), p. 93-99.

TANDON, A.; PATTERSON, N.; REICH, D. - Ancestry informative marker panels for African Americans based on subsets of commercially available SNP arrays. Genetic Epidemiology. ISSN 1098-2272. Vol. 35, n.º 1 (2011), p. 80-83.

TENG, X.; XIAO, H. - Perspectives of DNA microarray and next-generation DNA sequencing technologies. Science in China Series C: Life Sciences. ISSN 1006-9305. Vol. 52, n.º 1 (2009), p. 7-16.

THANGARAJ, K.; REDDY, A.; SINGH, L. - Is the amelogenin gene reliable for gender identification in forensic casework and prenatal diagnosis? International Journal of Legal Medicine. ISSN 0937-9827. Vol. 116, n.º 2 (2002), p. 121-123.

THE 1000 GENOMES PROJECT CONSORTIUM - A map of human genome variation from population-scale sequencing. Nature. ISSN 0028-0836. Vol. 467, n.º 7319 (2010), p. 1061-1073.

## Referências bibliográficas

THE GENETIC GENEALOGIST - DNAPrint Genomics Ceases Operations [em linha]. [Consult. 28-04-2011]. Disponível em WWW:<URL:<http://www.thegeneticgenealogist.com/2009/03/03/dnaprint-genomics-ceases-operations/>>.

THE INTERNATIONAL HAPMAP CONSORTIUM - A haplotype map of the human genome. Nature. Vol. 437, n.º 7063 (2005), p. 1299-1320.

THE UNIVERSITY OF UTAH - Learn.Genetics - Genetic Science Learning Center [em linha]. Salt Lake City. [Consult. 01-03-2011] Disponível em WWW:<URL:<http://learn.genetics.utah.edu/content/begin/traits/>>.

TISHKOFF, S. A.; KIDD, K. K. - Implications of biogeography of human populations for "race" and medicine. Nature Genetics Supplement. Vol. 36, n.º 11 (2004), p. s21-s27.

TULLY, G. [et al.] - Rapid Detection of Mitochondrial Sequence Polymorphisms Using Multiplex Solid-Phase Fluorescent Minisequencing. Genomics. ISSN 0888-7543. Vol. 34, n.º 1 (1996), p. 107-113.

VALENTE, C. [et al.] - Association between STRs from the X chromosome in a sample of Portuguese Gypsies. Forensic Science International: Genetics Supplement Series. ISSN 1875-1768. Vol. 2, n.º 1 (2009), p. 391-393.

VALENZUELA, R. K. [et al.] - Predicting Phenotype from Genotype: Normal Pigmentation\*. Journal of Forensic Sciences. ISSN 1556-4029. Vol. 55, n.º 2 (2010), p. 315-322.

VALVERDE, P. [et al.] - Variants of the melanocyte-stimulating hormone receptor gene are associated with red hair and fair skin in humans. Nat Genet. Vol. 11, n.º 3 (1995), p. 328-330.

VAN DAAL, A.; MANOHARAN, S.; NI, M.- DNA Identikit: Use of DNA Polymorphisms to Predict Offender Appearance: 18th International Symposium on Human Identification. Hollywood, CA: Promega, 2007.

VAN OVEN, M.; KAYSER, M. - Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. Human Mutation. ISSN 1098-1004. Vol. 30, n.º 2 (2009), p. E386-E394.

VIDEIRA, A. - Mitocôndrias. In: AZEVEDO, C. - Biologia Celular e Molecular. Lisboa: LIDEL, 1999. p. 245-154.

VISSCHER, P. M.; HILL, W. G.; WRAY, N. R. - Heritability in the genomics era — concepts and misconceptions. Nature Reviews Genetics. Vol. 9, n.º 4 (2008), p. 255-66.

VON WURMB-SCHWARK, N.; BOSINSKI, H.; RITZ-TIMME, S. - What do the X and Y chromosomes tell us about sex and gender in forensic case analysis? Journal of Forensic and Legal Medicine. ISSN 1752-928X. Vol. 14, n.º 1 (2007), p. 27-30.

WALSH, S. [et al.] - IrisPlex: A sensitive DNA tool for accurate prediction of blue and brown eye colour in the absence of ancestry information. Forensic science international. Genetics. ISSN 1872-4973. Vol. 5, n.º 3 (2011), p. 170-180.

- WANG, Z.; GERSTEIN, M.; SNYDER, M. - RNA-Seq: a revolutionary tool for transcriptomics. Nat Rev Genet. ISSN 1471-0056. Vol. 10, n.º 1 (2009), p. 57-63.
- WATKINS, W. S. [et al.] - Genetic Variation Among World Populations: Inferences From 100 Alu Insertion Polymorphisms. Genome Research. Vol. 13, n.º 7 (2003), p. 1607-1618.
- WEEDON, M. N. [et al.] - Genome-wide association analysis identifies 20 loci that influence adult height. Nat Genet. ISSN 1061-4036. Vol. 40, n.º 5 (2008), p. 575-583.
- WILLUWEIT, S.; ROEWER, L. - Y chromosome haplotype reference database (YHRD): Update. Forensic Science International: Genetics. ISSN 1872-4973. Vol. 1, n.º 2 (2007), p. 83-87.
- WILSON, A. C.; CANN, R. L. - The Recent African Genesis of Humans. Scientific American. (1992), p. 68-73.
- WITAS, H. W.; ZAWICKI, P. - Mitochondrial DNA and human evolution: A review. Anthropological Review. Vol. 67 (2004), p. 97-110.
- WU, D.; ZHANG, Y. - Different level of population differentiation among human genes. BMC Evolutionary Biology. ISSN 1471-2148. Vol. 11, n.º 1 (2011), p. 16.
- YANG, J. [et al.] - Common SNPs explain a large proportion of the heritability for human height. Nat Genet. ISSN 1061-4036. Vol. 42, n.º 7 (2010), p. 565-569.
- ZARAGOZA, M. V. [et al.] - Mitochondrial DNA Variant Discovery and Evaluation in Human Cardiomyopathies through Next-Generation Sequencing. PLoS ONE. Vol. 5, n.º 8 (2010), p. e12295.
- ZHANG, J. [et al.] - The impact of next-generation sequencing on genomics. Journal of Genetics and Genomics. ISSN 1673-8527. Vol. 38, n.º 3 (2011), p. 95-109.
- ZUBAKOV, D. [et al.] - Estimating human age from T-cell DNA rearrangements. Current biology : CB. ISSN 0960-9822. Vol. 20, n.º 22 (2010), p. R970-R971.



# ANEXOS

## Anexos

**Tabela A1 – Dados de 19 estudos relativos às bases genéticas dos traços de pigmentação (cor dos olhos, cor do cabelo e tom de pele) na qual se encontram discriminadas as associações estabelecidas entre diversos SNPs de vários genes e as referidas características bem como as populações nas quais essas relações foram estabelecidas (EUR – Europeia; ASI – Asiática; Várias – amostra incluía indivíduos com ancestralidade diversa). Na coluna denominada “Diferenças” encontram-se assinalados os trabalhos que reportam diferentes frequências alélicas em diferentes populações relativamente aos polimorfismos considerados).**

Gene	SNP	Fenótipos			Ancestralidade				Referências
		Cor dos Olhos	Cor do Cabelo	Tom de pele	EUR	ASI	Várias	Diferenças	
MC1R	MC1R_R		X		X				BRANICKI et al., 2011
			X	X	X				DUFFY et al., 2009
	MC1R_r		X		X				BRANICKI et al., 2011
	rs12931267		X		X				ERIKSSON et al., 2010
	rs885479	X		X			X		SPICHENOK et al., 2010
	rs1805006			X	X				DUFFY et al., 2009
	rs1805007			X			X		VALENZUELA et al., 2010
			X	X	X				DUFFY et al., 2009
	rs1805008		X	X	X				DUFFY et al., 2009
		X	X	X			X		VALENZUELA et al., 2010
			X		X				ERIKSSON et al., 2010
	rs1805009		X	*	X				SULEM et al., 2007
			X	X	X				DUFFY et al., 2009
			X		X				ERIKSSON et al., 2010
	rs2228478							X	NORTON et al., 2007
	rs2228479							X	NORTON et al., 2007
	rs3212346			X			X		VALENZUELA et al., 2010

\* - sardas; \*\* - capacidade de pigmentação facultativa.

Continua na página seguinte.



Tabela A1 continuação.

Gene	SNP	Fenótipos			Ancestralidade				Referências
		Cor dos Olhos	Cor do Cabelo	Tom de pele	EUR	ASI	Várias	Diferenças	
MC1R	rs3212355			X			X		VALENZUELA et al., 2010
	rs 3212357	X	X	X			X		VALENZUELA et al., 2010
HERC2	rs1129038	X			X				EIBERG et al., 2008
		X			X				LIU et al., 2010
		X			X				MENGEL-FROM et al., 2010
	rs11636232	X			X				EIBERG et al., 2008
		X			X				MENGEL-FROM et al., 2010
	rs12913832		X		X				BRANICKI et al., 2011
		X					X		WALSH et al., 2011
		X			X				MENGEL-FROM et al., 2010
			X	**	X				HAN et al., 2008
		X			X				LIU et al., 2010
			X				X		HANCOCK et al., 2011
		X			X				ERIKSSON et al., 2010
		X	X	X			X		VALENZUELA et al., 2010
		X	X	X	X				DUFFY et al., 2009
		X			X				EIBERG et al., 2008
		X		X			X		SPICHENOK et al., 2010
	rs2238289	X			X				EIBERG et al., 2008
		X			X				MENGEL-FROM et al., 2010

Continua na página seguinte.

Anexos

Tabela A1 continuação.

Gene	SNP	Fenótipos			Ancestralidade				Referências
		Cor dos Olhos	Cor do Cabelo	Tom de pele	EUR	ASI	Várias	Diferenças	
HERC2	rs3935591	X			X				EIBERG et al., 2008
	rs7170852	X			X				MENGEL-FROM et al., 2010
		X			X				EIBERG et al., 2008
HERC2/AP BA	rs2240203	X			X				EIBERG et al., 2008
	rs916977	X			X				MENGEL-FROM et al., 2010
		X			X				EIBERG et al., 2008
		X			X				KAYSER et al., 2008
IRF4	rs12203592	X			X				BRANICKI et al., 2011
		X							WALSH et al., 2011
		X	X	X e **	X				HAN et al., 2008
		X			X				LIU et al., 2010
		X		X			X		SPICHENOK et al., 2010
			X					X	STURM, 2009
			X	X	X				DUFFY et al., 2009
		X	X	*	X				ERIKSSON et al., 2010
	rs1540771		X					X	STURM, 2009
TYR	rs1042602		X		X				BRANICKI et al., 2011
								X	NORTON et al., 2007
				X				X	STURM, 2009

Continua na página seguinte.

Anexos

Tabela A1 continuação.

Gene	SNP	Fenótipos			Ancestralidade				Referências
		Cor dos Olhos	Cor do Cabelo	Tom de pele	EUR	ASI	Várias	Diferenças	
TYR	rs1042602			X		X			STOKOWSKI et al., 2007
				*	X				ERIKSSON et al., 2010
				*	X				SULEM et al., 2007
	rs10765198	X			X				DUFFY et al., 2009
	rs1126809	X		X	X				DUFFY et al., 2009
	rs1393350	X					X		WALSH et al., 2011
		X			X				ERIKSSON et al., 2010
		X			X				LIU et al., 2010
				**	X				SULEM et al., 2007
	rs1847134	X			X				ERIKSSON et al., 2010
TYRP1	rs1800422			X			X		STURM, 2009
	rs 683		X		X				BRANICKI et al., 2011
	rs10960751	X			X				ERIKSSON et al., 2010
	rs2733832	X	X	X			X		VALENZUELA et al., 2010
			X		X				DUFFY et al., 2009
	rs2762464							X	LAO et al., 2007
	rs1408799	X			X				SULEM et al., 2008
		X		X				X	STURM, 2009
			X		X				DUFFY et al., 2009

Continua na página seguinte.

Anexos

Tabela A1 continuação.

Gene	SNP	Fenótipos			Ancestralidade				Referências
		Cor dos Olhos	Cor do Cabelo	Tom de pele	EUR	ASI	Várias	Diferenças	
TYRP1	rs1325127	X			X				LIU et al., 2010
EXOC2	rs4959270		X		X				BRANICKI et al., 2011
	rs6918152		X		X				HAN et al., 2008
SLC24A4	rs2402130		X		X				BRANICKI et al., 2011
	rs12896399	X							WALSH et al., 2011
		X	X	**	X				HAN et al., 2008
			X		X				SULEM et al., 2007
		X	X					X	STURM, 2009
		X	X		X				DUFFY et al., 2009
		X			X				ERIKSSON et al., 2010
		X			X				LIU et al., 2010
SLC24A5	rs1426654							X	NORTON et al., 2007
				X				X	STURM, 2009
				X		X			STOKOWSKI et al., 2007
		X	X	X			X		VALENZUELA et al., 2010
				X			X		SPICHENOK et al., 2010
SLC45A2 (MATP)	rs26722			X				X	STURM, 2009
		X	X	X			X		VALENZUELA et al., 2010
		X			X				MENGEL-FROM et al., 2010

Continua na página seguinte.

Anexos

Tabela A1 continuação.

Gene	SNP	Fenótipos			Ancestralidade				Referências
		Cor dos Olhos	Cor do Cabelo	Tom de pele	EUR	ASI	Várias	Diferenças	
SLC45A2 (MATP)	rs28777		X		X				BRANICKI et al., 2011
			X	X e **	X				HAN et al, 2008
			X				X		HANCOCK et al., 2011
			X	X	X				DUFFY et al, 2009
	rs35391			**					HANCOCK et al., 2011
			X	X	X				DUFFY et al, 2009
	rs40132	X	X	X			X		VALENZUELA et al., 2010
	rs 16891982		X		X				BRANICKI et al., 2011
		X	X		X				ERIKSSON et al, 2010
		X		X			X		SPICHENOK et al, 2010
		X					X		WALSH et al., 2011
			X	X e **	X				HAN et al, 2008
								X	NORTON et al., 2007
								X	LAO et al, 2007
								X	STURM, 2009
				X		X			STOKOWSKI et al., 2007
		X			X				LIU et al, 2010
		X			X				MENGEL-FROM et al, 2010
		X	X	X			X		VALENZUELA et al., 2010

Continua na página seguinte.

Anexos

Tabela A1 continuação.

Gene	SNP	Fenótipos			Ancestralidade				Referências
		Cor dos Olhos	Cor do Cabelo	Tom de pele	EUR	ASI	Várias	Diferenças	
SLC45A2 (MATP)	rs16891982	X	X	X	X				DUFFY et al., 2009
	rs2287949	X		X			X		VALENZUELA et al., 2010
OCA2	rs1037208			X			X		VALENZUELA et al., 2010
	rs10852218			X			X		VALENZUELA et al., 2010
	rs11638265	X	X	X			X		VALENZUELA et al., 2010
	rs11855019	X			X				DUFFY et al., 2007
		X			X				KAYSER et al., 2008
	rs1800404	X		X					STURM, 2009
		X	X	X	X				VALENZUELA et al., 2010
								X	NORTON et al., 2007
	rs1800407		X		X				BRANICKI et al., 2011
		X					X		WALSH et al., 2011
								X	STURM, 2009
			X	X			X		VALENZUELA et al., 2010
		X			X				DUFFY et al., 2009
		X			X				MENGEL-FROM et al., 2010
		X			X				LIU et al., 2010
	rs1800410	X	X	X			X		VALENZUELA et al., 2010
	rs180411	X	X	X			X		VALENZUELA et al., 2010

Continua na página seguinte.

**Tabela A1 continuação.**

Gene	SNP	Fenótipos			Ancestralidade				Referências
		Cor dos Olhos	Cor do Cabelo	Tom de pele	EUR	ASI	Várias	Diferenças	
OCA2	rs1800414							X	LAO et al., 2007
								X	STURM, 2009
			X				X		VALENZUELA et al., 2010
	rs1900758	X	X	X			X		VALENZUELA et al., 2010
	rs6497268	X			X				KAYSER et al., 2008
		X			X				DUFFY et al., 2007
	rs728405	X			X				LIU et al., 2010
	rs7495174	X			X				DUFFY et al., 2007
		X			X				KAYSER et al., 2008
	rs749846	X	X				X		VALENZUELA et al., 2010
	rs4778241	X			X				EIBERG et al., 2008
	rs12913832	X						X	STURM, 2009
		X	X		X				ERIKSSON et al., 2010
	rs1448484							X	LAO et al., 2007
	rs1545397	X		X			X		SPICHENOK et al., 2010
	rs1667394	X	X		X				SULEM et al., 2007
		X			X				ERIKSSON et al., 2010
		X					X		HANCOCK et al., 2011
KITLG	rs12821256		X		X				BRANICKI et al., 2011

**Continua na página seguinte.**

Anexos

Tabela A1 continuação.

Gene	SNP	Fenótipos			Ancestralidade				Referências
		Cor dos Olhos	Cor do Cabelo	Tom de pele	EUR	ASI	Várias	Diferenças	
KITLG	rs 12821256		X		X				SULEM et al., 2007
			X					X	STURM, 2009
			X		X				ERIKSSON et al., 2010
	rs642742			X				X	STURM, 2009
ASIP	rs2378249		X		X				BRANICKI et al., 2011
	rs2424984	X	X	X			X		VALENZUELA et al., 2010
	rs4911442	X						X	STURM, 2009
			X	X	X				DUFFY et al., 2009
	rs 6058017							X	NORTON et al., 2007
								X	STURM, 2009
		X	X	X			X		VALENZUELA et al., 2010
	rs6119471	X		X			X		SPICHENOK et al., 2010
	Haplótipo (rs1015362, rs4911414)		X	* e **	X				SULEM et al., 2008
	rs1015362							X	STURM, 2009
DCT	rs4911414							X	STURM, 2009
	rs3782974							X	LAO et al., 2007
	rs1325611	X	X	X			X		VALENZUELA et al., 2010

Continua na página seguinte.



Anexos

**Tabela A1 continuação.**

Gene	SNP	Fenótipos			Ancestralidade				Referências
		Cor dos Olhos	Cor do Cabelo	Tom de pele	EUR	ASI	Várias	Diferenças	
DCT	rs1407995	X	X	X			X		VALENZUELA et al., 2010
CHR6	rs1540771			*	X				SULEM et al., 2007
TPCN2	rs35264875		X		X				SULEM et al., 2008
			X					X	STURM, 2009
	rs3829241		X		X				SULEM et al., 2008
			X					X	STURM, 2009
CYP4B1	rs157260			X			X		VALENZUELA et al., 2010
GPR143	rs3044	X	X	X			X		VALENZUELA et al., 2010
HPS4	rs1894704			X			X		VALENZUELA et al., 2010
	rs3752589	X					X		VALENZUELA et al., 2010
	rs3752590	X					X		VALENZUELA et al., 2010
MLPH	rs2292885		X	X			X		VALENZUELA et al., 2010
MYO5A	rs1724630			X			X		VALENZUELA et al., 2010
	rs752864			X			X		VALENZUELA et al., 2010
	rs2290332		X	X			X		VALENZUELA et al., 2010
MYO7A	rs2276289		X	X			X		VALENZUELA et al., 2010

**Continua na página seguinte.**

**Tabela A1 continuação.**

Gene	SNP	Fenótipos			Ancestralidade				Referências
		Cor dos Olhos	Cor do Cabelo	Tom de pele	EUR	ASI	Várias	Diferenças	
MYO7A	rs3737454		X	X					VALENZUELA et al., 2010
MYO18A	rs11080078	X					X		VALENZUELA et al., 2010
BNC2	rs2153271			*	X				ERIKSSON et al., 2010
LYST	rs3768056	X			X				LIU et al., 2010
NPLOC4	rs9894429	X			X				LIU et al., 2010
DSCR9	rs7277820	X			X				LIU et al., 2010

**Tabela A2– Estudos relativos a frequências alélicas de diversos marcadores que incluem amostras de populações Portuguesas.**

Material genético em Análise	Polimorfismos analisados		Referências
	Tipo	Conjunto de marcadores/regiões específicas (quando aplicável)	
Autossómico	STRs	AmpFI STR Profiler Plus e/ou PowerPlex 16 System	AMORIM et al., 2006, AMORIM et al., 2001, BRANCO et al., 2008c, FERNANDES et al., 2002b, GEADA et al., 2000, GUSMÃO et al., 2000, LOPES et al., 2009, PÉREZ-LEZAUN et al., 2000, PINHEIRO et al., 2005, VELOSA et al., 2002
		Outros conjuntos de marcadores	ABRANTES et al., 2003, ALVES et al., 2001, BRANCO et al., 2006a, BRANCO et al., 2008b, CRUZ et al., 2006, FERNANDES et al., 2002a, GUSMÃO et al., 2010, LAGOA et al., 2008, MANCO et al., 2007-2008, RIBEIRO et al., 2003
	SNPs	HLA	SPÍNOLA et al., 2002, SPÍNOLA et al., 2006
		SNPforID 52-plex ou subconjunto de marcadores	COSTA et al., 2008, PEREIRA et al., 2008
	Inserções Alu		BRANCO et al., 2006
X	STRs		CARVALHO et al., 2011, GUSMÃO et al., 2009, PEREIRA et al., 2007, SILVA et al., 2010a, VALENTE et al., 2009
	SNPs		PEREIRA et al., 2011
Y	STRs	Y-Plex (10 STRs) ou AmpFI STR® Y-filer (17 STRs)	ALVES et al., 2007, BENTO et al., 2008, BENTO et al., 2009b, FERNANDES et al., 2003, GUSMÃO et al., 2008b, PONTES et al., 2007
		Outros conjuntos de marcadores	CARVALHO et al., 2003, FERNANDES et al., 2001, GEADA et al., 2004, GONZÁLEZ-NEIRA et al., 2000, GUSMÃO et al., 2002, GUSMÃO et al., 2003b
	SNPs		CARVALHO et al., 2008, GONÇALVES et al., 2006, MONTIEL et al., 2005, ROSSER et al., 2000

**Continua na página seguinte.**

Tabela 2 continuação.

Material genético em Análise	Polimorfismos analisados		Referências
	Tipo	Conjunto de marcadores/regiões específicas (quando aplicável)	
Y	SNP e STRs		ADAMS et al., 2008, BELEZA et al., 2006, GUSMÃO et al., 2008b, GUSMÃO et al., 2003a, NOGUEIRO et al., 2010, PACHECO et al., 2005
	SNPs e inserções Alu		FLORES et al., 2004, PEREIRA et al., 2000a
mtDNA	SNPs	HVRI e/ou HVRII	LIMA et al., 2004, LIMA et al., 2006, PEREIRA et al., 2005, PEREIRA et al., 2003, PEREIRA et al., 2000b, PEREIRA et al., 2004
		HVI e RFLPs	BREHM et al., 2003, SANTOS et al., 2003
		Codificante	BENTO et al., 2009a, LIMA et al., 2008
Vários (marcadores de mais do que um tipo de material genético)			BRANCO et al., 2008a, PEREIRA et al., 2000c, PEREIRA et al., 2010

**Tabela A3 – Estudos relativos às frequências alélicas de diversas populações Portuguesas classificados de acordo com as populações amostradas.**

Populações/grupos Étnicos em estudo		Referências
Estudos que incluem amostras de uma única população	Norte 1	ABRANTES et al., 2003, ALVES et al., 2001, AMORIM et al., 2006, AMORIM et al., 2001, CARVALHO et al., 2011, FLORES et al., 2004, GONZÁLEZ-NEIRA et al., 2000, GUSMÃO et al., 2002, GUSMÃO et al., 2000, LAGOA et al., 2008, LIMA et al., 2004, LIMA et al., 2006, LIMA et al., 2008, PEREIRA et al., 2007, PÉREZ-LEZAUN et al., 2000, PINHEIRO et al., 2005, PONTES et al., 2007
	Centro 2	BENTO et al., 2009a, BENTO et al., 2008, BENTO et al., 2009b, CARVALHO et al., 2008, MANCO et al., 2007-2008
	Sul 3	COSTA et al., 2008, CRUZ et al., 2006, GEADA et al., 2004, PEREIRA et al., 2010, RIBEIRO et al., 2003
	Madeira 4	FERNANDES et al., 2002b, FERNANDES et al., 2001, SPÍNOLA et al., 2006
	Açores 5	BRANCO et al., 2008a, BRANCO et al., 2006a, BRANCO et al., 2008c, FERNANDES et al., 2003, MONTIEL et al., 2005, PACHECO et al., 2005, SANTOS et al., 2003, SILVA et al., 2010a, VELOSA et al., 2002
Estudos que incluem amostras de mais do que uma população	Norte e Sul 8	ADAMS et al., 2008, ROSSER et al., 2000
	Norte e Centro 16	GUSMÃO et al., 2009
	Centro e Açores 10	CARVALHO et al., 2003, LOPES et al., 2009
	Continente e Açores 13	BRANCO et al., 2008b, BRANCO et al., 2006b
	Madeira e Açores 9	BREHM et al., 2003

Continua na página seguinte.

Tabela A3 continuação.

Populações/grupos Étnicos em estudo		Referências
Estudos que incluem amostras de mais do que uma população	Norte, Centro e Sul 7	FERNANDES et al., 2002a, GUSMÃO et al., 2003b, PEREIRA et al., 2000a, PEREIRA et al., 2005, PEREIRA et al., 2000b, PEREIRA et al., 2004, PEREIRA et al., 2000c, PEREIRA et al., 2008, PEREIRA et al., 2011
	Norte, Centro, Sul e Madeira11	GUSMÃO et al., 2003a, SPÍNOLA et al., 2002
	Norte, Centro, Sul, Madeira e Açores12	GONÇALVES et al., 2006
	Norte, Portugal 14	PEREIRA et al., 2003
	Distritos e Regiões 15	BELEZA et al., 2006
	Geral/Indefinido	ALVES et al., 2007, GEADA et al., 2000
Estudos relativos a grupos étnicos	Ciganos	GUSMÃO et al., 2008a, GUSMÃO et al., 2008b, GUSMÃO et al., 2010, VALENTE et al., 2009
	Judeus	NOGUEIRO et al., 2010